3

REPORT RESUMES

ED 010 390 24
RELATIONSHIPS BETWEEN TEACHER CHARACTERISTICS AND STUDENT
BEHAVIOR--PART 111.
BY- WALLEN, NORMAN E.
UNIVERSITY OF UTAH, SALT LAKE CITY
REPORT NUMBER BR-5-1045 PUB DATE OCT 66
REPORT NUMBER CRP-2628
CONTRACT OEC-SAE-5-10-181
EDRS PRICE MF-$0.27 HC-$7.44 186P.

DESCRIPTORS- *TEACHER CHARACTERISTICS, *TEACHER EVALUATION, *STUDENT BEHAVIOR, *ELEMENTARY SCHOOL TEACHERS, PRIMARY GRADES, TEST RESULTS, *STUDENT TEACHER RELATIONSHIP, SALT LAKE CITY, UTAH

A SAMPLE OF 76 TEACHERS AND THEIR PUPILS, GRADES 1 AND 3, WAS USED IN THIS STUDY TO EXAMINE (1) RELATIONSHIPS BETWEEN SPECIFIC TEACHER MEASURES AND PUPIL (CLASS) CHANGE AND STATUS MEASURES, (2) CONSISTENCIES AMONG MEASURES OF TEACHER BEHAVIOR, (3) RELATIONSHIPS BETWEEN TEACHER BEHAVIOR AND TEACHER TEST DATA, AND (4) SPECIFIC TEACHER-PUPIL CLASSROOM INTERACTIONS. THE FOUR PARTS OF TEACHER BEHAVIOR STUDIED WERE CONTROL OF CLASS, AFFILIATION WITH THE STUDENT, STIMULATION OF PRESENTATION, AND ACADEMIC ACHIEVEMENT ORIENTATION. ON THE BASIS OF MEASUREMENTS OBTAINED USING THE STUDY SAMPLE AND ANALYSES OF RESULTING DATA, IT APPEARED THAT TEACHERS IN THE THIRD GRADE SHOULD BE STIMULATING AND INTELLECTUALLY EFFECTIVE AND, AT THE SAME TIME, WARM AND SUPPORTIVE PERSONS IF SUCH "DESIRABLE" OUTCOMES AS GAIN IN ACHIEVEMENT, LIKING FOR SCHOOL, LESS ANXIETY, AND INCREASE IN DIVERGENT THINKING ARE TO BE ACHIEVED. SUGGESTED QUALITIES OF GOOD FIRST-GRADE TEACHERS WERE ALSO STIMULATING AND INTELLECTUALLY EFFECTIVE CHARACTERISTICS, BUT SUCH OTHER CHARACTERISTICS AS OVERT AFFECTION AND STRONG PERSONALITY NEEDS FOR ACHIEVEMENT AND CONTROL WERE NEGATIVELY CORRELATED WITH "DESIRABLE" PUPIL OUTCOMES. CERTAIN PSYCHOLOGICAL TESTS USED IN THE STUDY OFFERED PROMISE IN THE PREDICTION OF SUCH TEACHER CHARACTERISTICS. OTHER RESEARCH EFFORTS CONDUCTED IN THE SUBJECT AREA, EITHER SINGLY OR JOINTLY BY THE AUTHOR OF THIS REPORT, WERE FUNDED, UNDER CONTRACTS OEC-444(8029), OEC-2-10-013, AND OEC-4-10-034. (JH)

ED010390

CRP 2628
5-1045

RELATIONSHIPS BETWEEN TEACHER CHARACTERISTICS AND STUDENT BEHAVIOR – PART III

NORMAN E. WALLEN

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY
UNIVERSITY OF UTAH
SALT LAKE CITY, UTAH 84112

OCTOBER 1966

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

The research reported herein is the Final Report of Contract No. SAE OE5-10-181 reported under the Cooperative Research Program of the Office of Education, U.S. Department of Health, Education and Welfare

RELATIONSHIPS BETWEEN TEACHER CHARACTERISTICS AND STUDENT BEHAVIOR--PART III

NORMAN E. WALLEN

Department of Educational Psychology
University of Utah
Salt Lake City, Utah 84112

October, 1966

The research reported herein is the Final Report of Contract No. SAE OE5-10-181 reported under the Cooperative Research Program of the Office of Education, U. S. Department of Health, Education, and Welfare.

ACKNOWLEDGEMENT

Appreciation is extended to the teachers, principals and district office personnel in the Granite School District whose cooperation made this study possible. Special appreciation is due Dr. Elmer V. Hartvigsen, Superintendent, and Mr. Ralf Riches, Elementary Supervisor.

Important contributions to this study were made by the following research assistants:

James Casagrande
Michael Hogben
Gary Jorgensen
Berthe Keith
Daniel Rhodes

Appreciation is extended to the Western Data Processing Center and to Shirley Schaefer and Thom Collins for services rendered in data processing.

Once again, it is a pleasure to acknowledge the always competent secretarial services of Carolyn Bennion.

ERRATA

Pages 3.24 and 3.28 delete or ignore columns immediately following the symbol 'f' and the symbol '%'

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENT ii

TABLE OF CONTENTS iii

CHAPTER

I. INTRODUCTION 1.01

II. PROCEDURES 2.01

Sample 2.01

Measures Obtained on the Pupils 2.03

Observational Measures of Teacher Behavior 2.09

Teacher Test Measures 2.19

Data Processing 2.20

III. MEASURES OF TEACHER CLASSROOM BEHAVIOR 3.01

Analysis of 'Observations' Data 3.01

Interrelationships Among Measures of Teacher Classroom . . Behavior 3.18

IV. CORRELATIONAL ANALYSES 4.01

Relationships Between Teacher Test Measures and Teacher Classroom Behavior 4.02

Relationships Between Teacher Classroom Behavior and Pupil Measures 4.19

Relationships Between Pupil Measures and Teacher Test Measures 4.43

Comparison of Grades (I and III) on Measures of Teacher Behavior and Teacher Tests 4.52

V. TEACHER-PUPIL INTERACTION 5.01

Analysis of Variance 5.01

CHAPTER Page

Grade Three 5.05

Results 5.09

VI. SUMMARY AND DISCUSSION 6.01

Adequacy of Observational Measures 6.02

Prediction of Teacher Behavior From Tests 6.03

Relationships Between Teacher Behavior and Pupil Behavior 6.05

Teacher Tests and Pupil Behavior 6.08

Teacher-Pupil Interaction 6.10

APPENDIX A PUPIL QUESTIONNAIRE (Keyed) A.02

APPENDIX B FIRST GRADE TEACHER MATRIX B.02

APPENDIX B THIRD GRADE TEACHER MATRIX B.10

APPENDIX C REGRESSION EQUATIONS USED IN OBTAINING PUPIL 'REGRESSED GAIN' SCORES C.02

SELECTED REFERENCES R.01

LIST OF TABLES

TABLE PAGE

3.01 AGREEMENT AMONG THE TWO MEMBERS OF EACH OBSERVATION TEAM ON INDEPENDENT SCORING OF SAME TEACHING SITUATION 3.05

3.02 ANOVA--OBSERVATIONS--CATEGORY 2, 'PRAISE AND ENCOURAGEMENT' 3.06

3.03 ANOVA--OBSERVATIONS--CATEGORY 3, 'MINIMAL REINFORCEMENT'. . 3.07

3.04 ANOVA--OBSERVATIONS--CATEGORY 4, 'ASKING QUESTIONS' 3.08

3.05 ANOVA--OBSERVATIONS--CATEGORY 5, 'PROBLEM STRUCTURING'. . . 3.09

3.06 ANOVA--OBSERVATIONS--CATEGORY 6, 'ACADEMIC CONTROL' 3.10

3.07 ANOVA--OBSERVATIONS--CATEGORY 7, 'PERSONAL CONTROL' 3.11

3.08 ANOVA--OBSERVATIONS--CATEGORY 8, 'HOSTILITY AND REPRIMANDS 3.12

3.09 COMPARISON OF TWO OBSERVER TEAMS ON MEANS IN MAJOR OBSERVATION CATEGORIES 3.13

3.10 CORRELATIONS BETWEEN 'JOINT' AND 'TOTAL' Q-SORTS GRADES I AND III 3.20

3.11 CORRELATIONS BETWEEN Q-SORTS (JOINT) AND RATINGS--GRADE I 3.21

3.12 CORRELATIONS BETWEEN Q-SORTS (JOINT) AND RATINGS--GRADE III 3.22

3.13 CORRELATIONS BETWEEN OBSERVATIONS (f and %) AND Q-SORTS (JOINT)--GRADE I 3.24

3.14 CORRELATIONS BETWEEN OBSERVATIONS (f and %) AND Q-SORTS (JOINT)--GRADE III 3.28

3.15 INTERCORRELATIONS AMONG Q-SORTS (JOINT)--GRADES I AND III 3.32

4.01 PREDICTED CORRELATIONS BETWEEN TPS CONTROL AND MEASURES OF TEACHER BEHAVIOR 4.03

4.02 NUMBER OF SIGNIFICANT CORRELATIONS OF TEACHER TEST MEASURES WITH TEACHER BEHAVIOR MEASURES 4.07

4.03 CORRELATIONS BETWEEN THE BEST PREDICTORS OF TEACHER BEHAVIOR AND THE TEACHER BEHAVIOR MEASURES 4.10

4.04 INTERCORRELATIONS AMONG THE BEST PREDICTORS OF TEACHER BEHAVIOR--GRADE I 4.13

TABLE PAGE

4.05 INTERCORRELATIONS AMONG THE BEST PREDICTORS OF TEACHER BEHAVIOR--GRADE III 4.20

4.06 CORRELATIONS BETWEEN CLASS ACHIEVEMENT AND MEASURES OF TEACHER BEHAVIOR 4.21

4.07 RELATIONSHIPS BETWEEN PUPIL QUESTIONNAIRE MEASURES OF ANXIETY AND TEACHER CHARACTERISTICS 4.24

4.08 'LIKING FOR SCHOOL' RELATED TO TEACHER BEHAVIOR--GRADE III 4.27

4.09 CORRELATIONS BETWEEN 'CIRCLES' TEST OF DIVERGENT THINKING AND MEASURES OF TEACHER BEHAVIOR--GRADE I 4.31

4.10 CORRELATIONS BETWEEN 'CIRCLES'TEST OF DIVERGENT THINKING AND MEASURES OF TEACHER BEHAVIOR--GRADE III 4.32

4.11 CORRELATIONS BETWEEN MEAN PEER EVALUATION BY CLASS AND MEASURES OF TEACHER BEHAVIOR--GRADE I 4.34

4.12 CORRELATIONS BETWEEN CLASS MEAN SOCIOMETRIC SCORES ON MEASURES OF TEACHER BEHAVIOR--GRADE III 4.36

4.13 CORRELATIONS BETWEEN MEASURES OF TEACHER BEHAVIOR AND RATINGS OF CLASS BEHAVIOR 4.39

4.14 RELATIONSHIPS BETWEEN TEACHER TEST MEASURES AND PUPIL ACHIEVEMENT 4.44

4.15 'LIKING FOR SCHOOL' AND TEACHER TESTS 4.47

4.16 RELATIONSHIPS BETWEEN TEACHER TEST VARIABLES AND PUPIL ANXIETY--GRADE ONE 4.49

4.17 TESTS PREDICTIVE OF OBSERVER RATINGS OF CLASS BEHAVIOR--GRADE III 4.50

4.18 COMPARISON OF GRADES I AND III ON MEASURES OF TEACHER BEHAVIOR 4.53

5.01 ANOVA--GRADE THREE GIRLS SELECTED ON READING VOCABULARY--GROUP A; DEPENDENT VARIABLES IS IN READING VOCABULARY . . 5.10

5.02 ANOVA--GRADE THREE GIRLS SELECTED ON READING VOCABULARY--GROUP B; DEPENDENT VARIABLE IS IN GAIN IN READING VOCABULARY. 5.11

5.03 MEANS--GRADE THREE GIRLS SELECTED ON READING VOCABULARY PRE-TEST--GROUP A; DEPENDENT VARIABLE IS GAIN IN READING VOCABULARY 5.11

5.04 MEANS--GRADE THREE GIRLS SELECTED ON READING VOCABULARY PRE-TEST--GROUP B; DEPENDENT VARIABLE IS GAIN IN READING VOCABULARY 5.12

LIST OF FIGURES

FIGURES PAGE

3.01 AGREEMENT OF FOUR OBSERVERS--TEACHER #1--GIRST GRADE . . 3.02

3.02 AGREEMENT OF FOUR OBSERVERS--TEACHER #2--SECOND GRADE . . 3.03

3.03 COMPARISON OF TWO TEACHERS ON 'OBSERVATIONS'--SCORES ARE MEANS BASED ON FOUR OBSERVERS 3.04

5.01 ANOVA DESIGN 5.06

CHAPTER I

INTRODUCTION

It is increasingly apparent that educational research must begin to establish relationships between certain environmental events to which students are exposed in school and consequent behavior of the students. A prerequisite to the study of such questions is the specification of differences in the environmental situations themselves. One major category of environmental differences is that of differences among teachers. A number of recent developments (for a review of this work see Wallen and Wodtke, 1963) have suggested meaningful dimensions of teacher difference and have begun to spell out in some detail techniques to be used in measuring these difference variables. We have been particularly interested in four dimensions of teacher behavior, time honored in educational theory but only recently subjected to rigorous experimental investigation. They are as follows:

1. Control--the degree to which the teacher controls the moment to moment behavior of her pupils.
2. Affiliation--the extent of warmth, support, and affection accorded pupils by the teacher.
3. Stimulation--the degree to which the teacher is stimulating, interesting, activity arousing to her pupils.
4. Achievement Orientation--the extent to which the teacher is focused on academic goals as opposed to focus on other educational goals such as interpersonal relations, morality,

contentment, etc.

The present study is the latest in a series of studies begun in 1958 by Dr. Robert M. W. Travers and the present author. The focus of the initial study was the attempt to spell out meaningful dimensions of teacher behavior in the classroom and the relating of such dimensions to paper and pencil teacher tests. The approaches used in the collection of teacher behavior in the classroom included analyses of verbatim classroom verbal behavior obtained on a time sampling procedure patterned after that of Withall, ratings made on the teachers after classroom observation by two observers, and a Q-sort procedure utilizing the same dimensions as the ratings but providing a slightly different technique intended to get at the same basic data; i.e., observer impressions of the teachers on a number of specified veriables including the ones of particular interest but also a number of variables gleaned from prior research. Paper and pencil measures administered to the teachers subsequent to the observations included a questionnaire measure of preference for classroom procedures and interaction with children--the Teacher Preference Schedule developed by Stern and Masling; the Utah Study of Behavior, a semi-projective device developed by Travers and Wallen; a preference questionnaire relating to educational values, attitudes, and philosophy; and a technique requesting the teacher to evaluate the relative merits of different approaches to the handling of specified classroom situations. These procedures were utilized with a total of 118 elementary school teachers in grades kindergarten through six constituting two separate samples one of N = 77 and the second of N = 41. The principal finding of this research was that one dimension of classroom behavior; extent of controlling behavior was fairly adequately assessed by the measures of classroom behavior and was predicted by a "control" score on the Teacher Preference Schedule; correlations ranging from .30 to .50 and replicated in the second sample. This study is summarized (Wallen, Travers, Reid, and Wodtke, Journal of Educational Psychology, 1963, Vol. 54, pp. 23-32) and is reported in detail in U. S. Office of Education Report of Contract No. 444(8029).

A subsequent study (Wallen and Wodtke, U. S. Office of Education Final Report, Contract No. 2-10-013, November, 1963) investigated the relationships between the various teacher measures previously collected and certain measures of pupil status and change for pupils of the same teachers in succeeding years. As a result of restricting the grade range to grades one through five and as a result of teacher turn-over, a total of 65 of the original of 118 teachers participated in this study. The measures obtained on the pupils consisted of pre- and post-test scores on the California Achievement tests--reading and arithmetic, and scores on several other measures obtained at various points in the school year. Tests included were selected tests of "creativity" from the Torrance Battery: a questionnaire including the Sarason Test Anxiety scale, the Medley and Cline 'Liking for School' Scale, and several other questionnaire scales devised as a part of the study and the Russell-Sage Social Relations Test--an indicator of group problem solving ability.

The mode of analysis consisted of obtaining regressed gain scores on the achievement measures for all pupils and relating these scores (as well as the status measures on pupils)--taken as mean class scores and hence as reflecting teacher behavior--to the measures of teacher behavior and teacher characteristics. This was done by means of Pearson correlation both within grades and as a composite across grades and through a series of factorial analyses setting up as main effects selected teacher characteristics and selected pupil characteristics and assessing the interaction as well as main effects as main effects in each analysis.

As a second part of this study, a sub-sample of teachers in grades two through five was selected so as to have the two teachers in each grade indicated as the most controlling teachers on the various measures and the two teachers within each grade indicated as the least controlling. For these teachers additional pupil measures were obtained in a subsequent academic year. Included here were pre- and post-test measures on selected 'creativity' tests, the Lorge-Thorndike group intelligence test, and observational measures of pupil classroom behavior. The conclusions of this project follow as reproduced from the original report

> It is clear that any conclusions arrived at as a result of a study such as this must be held quite tentatively. The first reason for this is that our sampling of teachers does not constitute a sample drawn randomly from teachers in general or even from teachers within the particular locale in which the study was conducted. We do have some confidence in generalizing results, at least within this geographical area, due to our attempt to obtain teachers from a number of schools and our evidence (from our earlier study) that the more important results pertaining to the prediction of teacher behavior were replicated with both subsamples. Nevertheless, this caution should be kept in mind.
>
> Secondly, most of the hypotheses which we originally developed and which were directly tested were not supported. Thus, we have in large measure resorted to an empirical examination of the many relationships possible within our matrix of variables. Although we have attempted to guard against capitalization on chance, there remains the possibility that some of our findings reflect nothing more than chance fluctuations. This problem is particularly evident when the relationships are examined within a particular grade with the smaller N which this

necessitates. Unfortunately within-grade analysis became imperative in this study due to the finding that one of our initial assumptions, i.e., that any relationships obtained between teacher characteristics and student characteristics would take essentially the same form across all five grades is, in many instances, apparently an incorrect assumption, although one cannot be sure whether it is the relationship between the fundamental variables themselves which differs or whether certain aspects of the measurement procedures may have differed from grade to grade.

Nevertheless, it is possible to place our findings along a continuum of tentativeness and we shall first list those results which seem to us to have considerable support within our data both in terms of consistencies across the five grades and in terms of the consistency of relationships among the variables themselves.

1. Achievement gain in Reading Vocabulary appears to be positively correlated with the extent to which the teacher is viewed as a stimulating teacher by observers. This relationship seems to exist within all five grades.

2. 'Liking for School' appears to be related to a teacher orientation which is less achievement oriented as viewed by observers. It is probably to be expected that the teacher who makes the greatest achievement and academic demands on the student will engender the greater amount of frustration on the part of the poor students in particular, and hence, a somewhat less favorable attitude toward her and school in general.

3. 'Liking for School' seems to be positively related to the degree of warmth and permissiveness of the teacher in the upper grades. In the lower grades, and particularly in grade one, this relationship is possibly reversed, although this finding is considered very tentative due to the complications in interpretation of the measure of permissiveness encountered in this grade.

4. The extent to which the teacher is viewed by observers as achievement oriented seems to have a positive affect upon the ability of the class to plan effectively in a group problem solving task, but also appears to foster a breakdown in group relationships in the actual operations phase of the problem though this did not appear until the second of two problems was encountered.

5. Supportive behavior on the part of the teacher appears to foster a more friendly group interaction during the operations phase of group problem solving activity.

6. Supportive behavior on the part of the teacher appears to be negatively correlated with the extent of test anxiety within all five grades.

7. There appears to be a fair degree of correspondence between observer and pupil perception of the teacher in terms of extent of affiliation behavior beginning at about the third grade. In the first two grades this correspondence is extremely poor.

8. Speaking only of achievement gain in the three areas tested, it appears that some teachers do a better job of teaching reading, across all grades than others. Within the first grade, there is the suggestion that the teacher who fosters greater gain in Reading Comprehension also fosters greater gain in Arithmetic. This trend was also observed in the upper grades but may (in the upper grades) be attributable to differences in class intelligence.

9. Coefficients of stability for the Torrance creativity tests over a six-month period show greater stability in the higher grades and higher stability for the verbal than for the non-verbal measures. In grades four and five, the 'total' score in each area (verbal and non-verbal) provides stability coefficients of the magnitude of .60 to .75. In the lower grades (two and three) the non-verbal 'total' stability values are .46 and .34.

Next we may list a series of findings which we regard as even more tentative than those listed above.

1. It appears that the teacher who is excessively affiliative or affectionate in the first grade and who, in addition, has a strong need for affiliation is likely to be less well liked and to engender less of an achievement gain than the less affiliation oriented teacher, although hostility in the first grade on the part of the teacher appears undesirable. It appears that verbal supportive behavior, praise, and encouragement on the part of the teacher is quite important in the first grade, fostering both liking for school and gain in achievement. It also appears, however, that in the first grade the student may have need of considerable structure within the school environment such that the more controlled classroom and the controlling, somewhat aloof teacher may provide an environment within this grade which is more comfortable and more conducive to achievement for the first grade student. It would appear that the first grade teacher should provide a very well ordered, well organized classroom which does not permit much pupil decision but which is at the same time comforting and encouraging but not overtly affectionate.

2. There is the suggestion of an interaction between the teacher characteristic of control and the pupil characteristic of dependency in that within the first grade on the reading vocabulary measure the children who were high in dependency need, as measured by our questionnaire, did better with the more controlling teachers, whereas, the children who were low in dependency need did better with the less controlling teachers.

One analysis suggests the same effect to be true for arithmetic gain. There is the further suggestion that this relationship is reversed by the fifth grade in that the less dependent children achieved less achievement gain in arithmetic with the less controlling teachers.

3. With respect to changes in creativity, at least as measured by the Torrance tests, there is some evidence that the more permissive teacher in the upper grades tends to foster a greater change in the direction of more originality of thinking than does the more controlling teacher.

4. There is the suggestion that pupils, at least in the upper grades, initiate more verbal exchange with the less controlling teacher.

5. On several fairly reliable indexes of pupil classroom behavior, there was no indication of differing behavior on the part of 'high' or 'low' creative pupils.

In addition to the generalizations just discussed, it seems worth recording the overall impression which we are left with; an impression which, it should be recognized, would not necessarily be shared by other individuals working with the same data since it does not have clear-cut support from the data. Nevertheless, the impression is one of the importance of a developmental point of view with regard to teacher behavior across the first five grades. We have the distinct impression from our data that the typical first grader, being somewhat unsure of himself and new to the situation, is both more comfortable and achieves better given a situation which is quite structured, quite controlled by the teacher while at the same time being rather supportive and encouraging but without overt affection on the part of the teacher. Data which tend to support this are the preceding discussion of the results within grade one as well as the finding that this type of description seems to hold particularly true for the more dependent child. In about the second grade it is our impression that this pattern tends to change and that by the upper grades the general desirability of encouragement still exists but the effect of control shifts such that a greater degree of permissiveness has the more desirable effects in terms of both 'liking for school' and achievement gain.

Further, the totality of results strongly suggest that relationships between teacher characteristics and pupil behavior may be very different in the first one or two grades than for the succeeding grades. Since almost all research has been done using grades four and above (probably largely for reasons of convenience which we can well appreciate), we suggest that: 1. generalization below these grades is very questionable and 2. studies in the first grade seem imperative.

A subsequent study (Wallen, U. S. Office of Education, Project No. OE-4-10-034, Final Report of August, 1964) conducted further analyses of these same data primarily to clarify some of the tentative interpretations just discussed, but without much success. In addition, the data were analyzed from the point of view shown to be profitably by Heil, et al., but we were unable to replicate the Heil findings. Various interpretations of this are presented in that report.

The present study is an attempt to pursue some of the interesting prior findings to a more conclusive outcome by utilizing an entirely new sample of teachers and pupils; collecting all the data within one academic year and utilizing teacher samples of size 40 in each of two selected grades--grades one and three. The specific procedures followed are contained in the following chapter.

CHAPTER II

PROCEDURES

Sample

It was planned to utilize a sample of 40 female teachers in each of grades one and three. All teachers were within one large metropolitan school district in the urban Salt Lake City area. It was planned to obtain a random sampling of teachers within the district subject only to the limitations of: 1. An attempt to stratify schools by socioeconomic status; 2. Use of several teachers within the same school to obviate excessive transportation problems; and 3. Use of no teacher in her first year of teaching. Thus, it was hoped that a minimum of four teachers could be obtained within each school, necessitating that approximately 20 different elementary schools be utilized. The selection of schools was done by the elementary supervisor in the district in question, keeping in mind the restrictions just discussed. It is our feeling that this procedure was accomplished relatively well subject to certain biases which are likely inevitable in this type of research. First, the supervisor and the district needed to approach schools where it was felt likely that a willingness to participate would be found. Some of the schools were participating in other projects of various kinds, and hence some negative selection may have occurred on this variable. In addition, it is likely that those principals whose attitudes were assessed as most cooperative would have been selected. It is our impression, however, that our final group of teachers does not depart radically from a stratified random sample of teachers within this school district.

The cooperation of teachers and principals was solicited by personnel from the school district office. Throughout the study most of the teachers participated graciously; a few enthusiastically. There were, however, the seemingly inevitable questions as to the merit of particular instruments and the purpose of the study as well as errors in scheduling and occasional resistance to being observed. It has been our experience that such problems are rather easily handled once the teachers become acquainted with the project staff and provided they are given answers to their queries and reassurance as to the anonymous status of teachers in treatment of data.

The study began with 40 first-grade teachers and 42 third-grade teachers. Due to illness, organizational changes, and teachers leaving the system, the final sample of teachers consisted of 36 first-grade and 40 third-grade teachers. Observational and test-type data were obtained for all of these teachers and all of their pupils with the exception of pupil absences from class on the particular days on which certain instruments were given. The first grades attended school half-days only--a session of 3½ hours per day. In the first grade, the class size hovered very closely around 22 and for the third grade classes around 32. A total of 16 schools were involved and all teachers of the first and third grades in these schools participated in the study with the exception of one school in which the third grade teachers did not wish to participate and a second school in which the third grade teachers were first year teachers and hence not included.

Measures Obtained on the Pupils

The method of data collection was as follows: The total sample of schools was divided into two sub-samples each including approximately half of the teachers at each grade. Two research assistants were responsible for the data collection in each of these sub-samples of schools. Listed below are the measures obtained and the approximate dates during which the data were collected. It will be noted that a particular instrument was administered to the pupils of all teachers within approximately a two-week period. The time interval between tests varied from 5½ months to 7½ months.

Pre-Test

Test	Dates
First Grade	
California Achievement-Reading Vocabulary--Lower Primary	10/20/64-10/23/64
Circles Test from Torrance Battery	10/20/64-10/23/64
Questionnaire	10/30/64-11/5/64
Third Grade	
California Achievement-Reading Vocabulary--Elementary Level	10/27/64-11/4/65
Circles Test from Torrance Battery	10/27/64-11/4/65
Questionnaire	11/9/64-11/12/64
Barron-Welsh	11/9/64-11/20/64

Post-Test

First Grade	
California Achievement Test Reading-Vocabulary and Comprehension--Lower Primary	4/5/65-4/9/65
California Achievement Test Arithmetic Computation--Lower Primary	4/5/65-4/14/65
Circles Test from Torrance Battery	5/13/65-5/26-65

Questionnaire	5/13/65/-5/26/65
Sociometric	4/27/65-5/12/65
Third Grade	
California Achievement Test Reading-Vocabulary and Comprehension--Elementary Level	4/8/65-4/31/65
California Achievement Test Arithmetic Computation--Elementary Level	4/21/65-4/30/65
Circles Test from Torrance Battery	5/13/65-4/30/65
Questionnaire	5/13/65-5/25/65
Sociometric	5/21/65-6/3/65
Barron-Welsh	4/27/65-5/19/65

The achievement measure used was the California Achievement Test. In the fall, or pre-testing, only the reading vocabulary subtest was administered since our previous research had shown that it correlated very nearly ashighly with end-of-year scores as fall scores on these same tests. Thus, in obtaining the regressed gain scores, the reading vocabulary pre-test functioned as well as the pre-test on the comprehension and arithmetic measures themselves and required less testing. Since our previous research had suggested that the difficulty level of these tests as applied to children in this geographic area was somewhat inappropriate, the levels of the test used were not in all instances those suggested by the test publisher. The Lower Primary level was utilized in the first grade and the Elementary level was utilized in the third grade.

A second measure utilized was the Circles Test from the Torrance Creativity Battery. It was considered of interest to study possible changes on a measure of divergent thinking (such as this is reported to be) as a function of teacher behavior. Since our previous research had suggested that the Circles Test was the most reliable of the non-verbal tasks, and hence this test was administered both in the fall and in the spring for both grades. The test requires the respondent to construct figures utilizing a page composed of a large number of circles. His task is to make as many different and interesting objects as he can within the time limit. Scoring procedures outlined by Yamamoto were followed. In addition, since the scorers felt the scoring gave excessive weight to sheer fluency, a second score (Creativity Rating) was used.

The creativity ratings were done by three raters. A seven-point rating scale was developed and applied to protocols which had been collected during the previous project. The raters first worked independently and later jointly on the same protocols in order to resolve differences and establish the criteria for rating. A sample of 60 protocols rated independently by the three judges yielded intercorrelations all above .85. Because the procedure was fairly easy to do and could be accomplished in a short period of time, it was felt that the whole group of raters should do every rating jointly, resolving differences as they arose. Artistic ability may have entered into the ratings, but this was not one of the criteria, and efforts were made to reduce the effect of this bias by stressing the following criteria:

1. An unusual or relatively infrequent response. (This was given high priority.)
2. The sheer number of unusual or relatively infrequent responses was an important factor. (If many unusual or relatively infrequent responses were given, this tended to increase the rating.)
3. The use of multiple figures tended to raise the score. (The

use of two or more circles in an integrated fashion was thought to be a sign of greater flexibility and integrative powers.)

4. The sheer number of non-duplicated responses. (This was of lesser importance because the responses tended to be duplicated from person to person and class to class.)

Another measure utilized was a questionnaire consisting of items to which the student responded yes or no (Appendix A). All questions were read aloud to the class, and they indicated their responses on answer sheets.

Included in the questionnaire were the Sarason Test Anxiety Scale, Medley and Kline 'Liking for School,'* both of which have been utilized by ourselves and others in prior research, and six other scales utilized by us in previous studies consisting of items intended to get at the following variables:

1. Need achievement
2. Dependency need
3. Affiliation need
4. Perception of teacher as controlling
5. Perception of teacher as affiliative
6. Perception of teacher as achievement oriented

In addition to these measures, the Barron-Welsh Art Scale was administered fall and spring in the third grade only. This commercially published testhas some demonstrated validity in distinguishing highly creative groups of adults. It has been utilized to a limited extent with children, to our knowledge never as low as the third grade, but was considered of sufficient promise to be included in this study. The task set the child is to indicate liking or dislike for each of a series of 80 designs.

---

*The nature of these questions is such that we prefer to view it as a 'liking for teacher' measure and often refer to it as such.

In addition to these measures, sociometric data were obtained in the spring only for both first and third grades. The technique utilized was patterned very closely after the Ohio Social Acceptance Scale. Names of all of the members of the class were printed on sheets which were passed out to the pupils. They were to indicate placement of each individual by marking the appropriate place on a continuous scale. In the first grade, this was done by utilizing the bridge drawing commonly used in first grade, which is simply a curved line; the child places an X on the line from most to least. The continuous line was subsequently divided into five categories and the position of the check converted to numerical score. In the third grade, the same basic procedure was utilized with the exception that the bridge was no longer used and the children were simply instructed to place checks in the appropriate categories (1 through 5). In grade one, the dimension rated was simply degree of liking. In grade three a more elaborate set of dimensions was used as indicated in the following directions.

Directions for Sociometric Device--Third Grade

Passing out booklets: You are about to get a booklet which is stapled together--please do not separate any of the pages from the booklet.

1. Aggression.--We are going to play some interesting games with these booklets. Each person's name is on one of the pages. Each of you will rate all your classmates and yourself on five characteristics, or ways you usually act. This is how it will be done: (draw boxes) The first characteristic we shall rate is willingness to get into fights.

For the people who like to start fights, argue, and boss others around, you will put an "X" in the box beside their name under the top box marked "most"--that means they most like to fight (demonstrate). For the people who least like to fight and argue, we will put an "X" in the box next to their name under the top box marked "least." Alot of people will be marked in the medium box since they occasionally fight and argue, but not nearly as much as those for whom you check most. Quite a few people will fall in the between boxes; under the one between medium and most go those who like to fight and argue, but not as much as those you rate most. Under the other between box you will mark those people who don't like to fight or argue, but once in a while do this. Remember: I am not saying that any of these characteristics are good or bad--I merely want you to describe the rest of your classmates and yourself. Put only one mark next to each person's name in the box where you think he or she belongs. When I call your name, circle it and rate yourself like you have for everyone else. Also, most people will go into the Between and Medium boxes, and only a small number go into each end box.

(Note: Emphasize with dimensions 1, 2, and 4--not good or bad, just want a description; emphasize with dimensions 3 and 5--most people will go into the between or medium boxes; after dimension 1: repeat: circle your own name when you rate yourself, and don't skip any names.)

2. Dependency most.--The children who like to be told again and again exactly how to do things when the teacher starts a new activity. They ask a lot of questions about how things should be done and if they are doing them right. They are very unwilling to do things on their own. In the least category go people who like to do things on their own and try out new jobs and assignments all by themselves. In

the medium box go people who want to ask some questions about how to do things, but usually wait for a while to see if they can do them on their own.

3. Achievement most.--Mark the children who want to get the highest grades on all their work. Not necessarily that they do get them, but they want to very badly. In the least boxes go those people who are not concerned with the marks they get. Not that they necessarily do poorly in their work, but they care very little about it.

4. Affiliation most.--These are the friendliest people in the class. Not necessarily your best friend, but rather, the friendliest people. In the least boxes go those people who keep to themselves and aren't very friendly. This does not mean that they aren't nice, but that they are simply not very friendly.

5. Test Anxiety most.--The people who get very upset, nervous and uncomfortable when the teacher says she is going to give a test to find out how much you know. Least: The people who simply take out their pencils and papers when the teacher announces a test, without getting nervous or upset. Tests don't bother them at all.

Scores: Each pupil score was the mean rating received in each category.

Observational Measures of Teacher Behavior

Having had considerable experience with various procedures for obtaining data on classroom behavior of teachers, including the modified Withall technique applied in the study with Travers, a limited tryout of the OSCAR technique in the study with Wodtke, and after some tryout on the present study, a modification of Flanders' approach was utilized.

This approach requires the observer to sit in the classroom for thirty-minute periods and check each item of behavior on the part of the teacher which falls into the classification system described below. Our approach differs from Flanders in two basic ways. First, we have found it necessary to change a number of the categories since tryout showed that some of the Flanders' categories are quite inapplicable in the first grade and of questionable appropriateness in the third grade. Hence, we have made modifications in the categories to be observed. Secondly, in the Flanders' technique, one attempts to record almost continuously, obtaining as many as 600 notations in a thirty-minute period and collects these in sequence noting down numbers corresponding to the categories, thus permitting the construction of Flanders' interaction matrix which provides not only an index of the frequency of various categories of behavior but a sequencing of such categories. In the present study this sequencing was considered to be of insufficient value to justify the additional labor of subsequently categorizing and grouping these figures. Hence, the approach followed was simply to check the behavior in the various categories with no attempt at sequencing. Further, the focus was limited to teach behavior only. It was intended that all teacher verbal behavior be recorded, as well as those non-verbal behaviors fitting the defined categories. The unit or scoring was considered to be an information segment. Thus, a long informative statement might receive only one scored 5. The observer would follow grammatical structure in deciding whether to score a second 5. If two distinct sentences were stated, two entries would be made.

A considerable amount of time was spent in the development and tryout of these categories both at the laboratory school on campus and in a school utilized specifically for this purpose which was not included in the major data collection. The basic approach to the training of observers included initial sitting in classrooms on the part of all of the four research assistants and the project director and attempting to classify behavior, followed immediately by discussion of specific teacher behaviors and the general trend of scoring obtained. Interspersed with this was the listening to tapes of teacher-pupil interaction which were on hand as a result of prior research and the scoring of sequences of teacher verbal statements followed by discussion. Following this the procedures were applied to teachers with decreasing amounts of discussion and hence greater objectivity and independence of scoring. Prior to the collection of data on the teachers who constituted the major sample, independent scoring by the five observers gave agreement which was considered satisfactory, i.e., different observers obtained very similar frequencies in the same category.

Following this training procedure, the observational technique* was applied to the major sample of teachers in the following manner. Each pair of research assistants observed in the classrooms of the teachers whose children they had previously tested and with whom they were now acquainted. It was initially planned that each assistant would visit the classroom for one warm-up period without collecting data simply to accustom the teacher and pupils to his or her presence. After several of these visits, however, it was our impression that this 'warm-up' was unnecessary and data collection began on the first such visit. Each of the observers visited each teacher for five one-half hour periods. A total of nine visits to each teacher resulted, since one of these visits was accomplished jointly by the observers for the purposes of obtaining additional data as to how well the observers were agreeing and also as a way of helping the observers keep on their toes with regard to the data collection. The observations were collected during the period from January 15 through March 31. Observations in each school were collected during a two-week interval.

---

*Referred to hereafter as 'Observations.'

Categories Used in Observation of Teacher Behavior

1. Acknowledges student's raised hand.--Self explanatory.
2. Praise and encouragement.--Includes generally supportive behavior and behavior viewed as positively reinforcing. May include value judgments on the part of the teacher if they are positive. Examples: "Very good; that's right; fine."

2a. Non-verbal affiliation.--Physical contact with the student such as putting arm around student.

3. Minimal reinforcement.--Includes positive feedback to student where it is the observer's impression that this is not a strong reinforcement; for example, "Uh huh; go on; okay; right; etc." The teacher is in a sense reinforcing the pupil but at a minimal level and with essentially no emotional overtones. In most cases the student does not respond overtly.
4. Asking questions with the intent that the student answer.--Example: "How many pennies make a dollar? What was the point of the story, etc.?"
5. Explaining or problem structuring.--Includes helping with words while reading, clarifying material, directive statements closely tied to content. Also includes correcting errors.

6. Academic control.--Teacher directs the students to perform certain actions clearly related to academic learning. This includes a rhetorical question. Examples: "Open your books; come to my desk for help; read page 13." Also includes calling on a student who does not have his hand raised or volunteers in any fashion. Example: "Speak louder and with expression; group two come to the front; open your books; watch me; George, will you read page 32; look at the board; be sure you have the right page."

7. Personal control.--The teacher directs the students to perform or to stop certain actions related to personal rather than academic behavior. Examples: Teacher rearranges pupils chairs; "Lay your pencil down when you are through; put your hands on your head when you are finished; go to your seat; sit up straight, Johnny, you don't want to do that do you; Johnny, do you need some help (in the context of breaking up an argument)." Includes statements intended to influerce the student's behavior when disapproved of by the teacher but which are not strong enough to be viewed as a reprimand or as a hostile comment.

7m. Moralizing by teacher.--Example: "Don't do that, you wouldn't like it if Johnny did that to you."

8. Hostility and reprimands.--Any teacher behavior which is definitely antagonistic toward students. Examples: "Be quiet (sternly);" physically striking the child; shaking the head from side to side in a very scolding manner.

8x. Ignores child's behavior when a child is attempting to get the teacher's attention.

In using the observation, it is possible for a given teacher behavior to receive more than one category. For example, if the teacher compliments one student as an attempt to control others, this would be scored for both a 7 (personal control) and then either a 2 or a 3 for the reinforcement of another pupil, depending upon the strength of the reinforcement. Example: I like the way Jean sits up (7.2).

Two major types of scores were derived. The first is intended to measure differences among teachers with regard to the frequency of occurrence of behavior within each category ('f' in subsequent tables). To this end, the total frequency (across five visits) was obtained for a given category for each teacher--for each of the two observers assigned to her. Next the distribution of such scores for each observer (within each grade) was divided into nine parts. Since the distributionsare, for the most part, unimodal and roughly normal, scores expressed in these units correspond roughly to 'stainines.' Assigning a 'stanine' based on each observer is intended to rule out consistent tendencies for some observers to record higher frequencies in certain categories. Further, basing the stanines on separate distributions for the two grades is intended to rule out systematic 'between grades' differences. Next the 'stanine' for the two observers were summed to provide the final score--based on a total of ten observer visits. In the results section this score is referred to as the frequency (f) score.

The second score is intended to measure differences among teachers with regard to the proportion (% in subsequent tables) of observed behavior falling within each category, since the total frequency of categorizable behavior differs from teacher to teacher. Thus, the total frequency (both observers) within each category was expressed as a proportion of the total frequency across all categories (both observers).

In addition, analyses of variance were computed using the original frequency for each observation, permitting assessment of the contribution to total variance made by teachers, observers, schools, and grades.

Ratings and Q-Sorts

In addition to the observational measures just discussed, additional measures of teacher behavior were obtained as follows: Subsequent to each observational period, each observer rated the teacher on a seven-point scale with regard to the following characteristics as manifested during that particular period of time:

Teacher Behavior:

1. Permissive vs. Controlling
2. Dull vs. Stimulating
3. Disparaging vs. Not Disparaging
4. Supportive vs. Not Supportive
5. Anxious vs. Confident
6. Aloof vs. Affiliative
7. Intent vs. Relaxed
8. Smiling vs. Sour

In addition, the class (pupils) was rated on three variables:

1. Orderly vs. Chaotic
2. Happy vs. Unhappy
3. Independent vs. Dependent (on teachers or one another)

After the observations on all teachers were completed each of the observers independently Q-sorted the sample of teachers whom he had observed. Following this the two observers working together as a team jointly Q-sorted.

In Q-sorting, each pair of observers wrote the names of the teachers they had observed on cards. With the teachers' names on the cards, teachers were placed in one of seven piles or scores with respect to each of the variables. For example, on variable one--controlling behavior--the most controlling teachers were put in pile one, the moderately controlling teachers in pile four, and the least controlling teachers in pile seven.*

Of the approximately 40 teachers for each pair of observers (the first and third grades were done together), the different piles or scores were distributed approximately as follows:

Score received:	1	2	3	4	5	6	7	Total
No. of teachers:	3	5	7	10	7	5	3	40
Per cent of teachers:	7.5	12.5	17.5	25	17.5	12.5	7.5	100.0

The dimensions or variables for the teachers were as follows:

1. 1 High Controlling vs. Low Controlling 7
2. 1 Warm and Affiliative vs. Cool and Aloof 7
3. 1 Punitive vs. Non-punitive 7
4. 1 Confident vs. Anxious and Uncertain 7
5. 1 Supportive vs. Non-supportive 7
6. 1 Dynamic and Stimulating vs. Dull 7
7. 1 Achievement Oriented vs. Non-achievement Oriented 7
8. 1 Much Physical Contact vs. Little Physical Contact 7
9. 1 Intellectually Effective vs. Intellectually Ineffective
10. 1 Disparaging vs. Non-disparaging 7

---

*Note: In the results section, each variable is discussed as defined by the '1' end of the scale. Signs of correlation coefficients are reversed accordingly.

In Q-sorting many of the dimensions were similar to those in the observation categories and ratings; however, in Q-sorting the observers were allowed more subjectivity. Definitions of the Q-sort variables were as follows:

1. Controlling.--In the observation categories the sixes and sevens were controlling categories; on the 'ratings' the instructions were that the teacher who made many "attempts to control, despite her success" was rated high on the control variables. In Q-sorting, both the teacher who was constantly making controlling statements, with some success, and the teacher whose students were very restrained in her presence despite a lesser number of controlling statements, were placed high (1) in this category. In the other direction (7) were those teachers who both attempted less and achieved less control.

2. Warm and Affiliative vs. Cool and Aloof.--On this variable, the teacher who was judged to demonstrate genuine affiliation was rated as higher than the teacher with a permanent smile or the one who gave many reinforcements. On the low end of the scale (7) were both the more aloof impersonal teachers and the hostile teachers.

3. Punitive.--The punitive teacher (1 high, 7 low) is rather self-explanatory. She is one who excessively punishes, physically and verbally, her class or specific individuals.

4. Confident vs. Anxious and Uncertain.--The teacher who appeared to be confident and assured in her role as a teacher rated high (1) on this variable. The anxious, nervous teacher was rated low (7).

5. Supportive.--The supportive teacher may or may not have been the affiliative teacher. The teachers who rated high (1) on this variable gave more encouragement and praise--reinforcement--than those who rated

low (7).

6. Dynamic and Stimulating.--The stimulating teacher (1) is the one who, in the observer's opinion, was most interesting and/or dynamic in her role as a teacher. The teachers who rated low (7) were considered less stimulating.

7. Achievement Oriented.--Some teachers appeared more concerned that the students learn the material; others appeared more interested (particularly in the first grade) that the students enjoy themselves; or at least didn't demonstrate to the observers as much concern about how much their students assimilated. Those who were most concerned about academic learning rated high (1) on this dimension. Those least concerned, low (7).

8. Physical Contact.--The teacher who tended to hug and fondle rated high (1) on this dimension. The physically aloof teacher rated low (7).

9. Intellectually Effective.--The intellectually effective teacher (1) was the one who appeared to be able to explain concepts clearly and such that the students seemed to be gaining understanding. She had facility with her material and enough background to answer intelligently her children's many questions. The teacher who rated low (7) impressed the observers less in these respects.

10. Disparaging.--The highly disparaging teacher (1) was continually criticizing her class or students; whereas the minimally disparaging teacher was much less critical, though not necessarily highly supportive.

In addition, one other variable, based on our prior work was included. It is a combination of 'Control' and 'Warmth' scales and is labeled 'Factor I.' Low scores indicate: 'Warm, Low Control;' high scores indicate 'Aloof, Much Control.'

Teacher Test Measures

Following the collection of all other data, i.e., pre- and post-pupil measures and the observation of teacher behavior, each teacher filled out a number of devices in order as follows:

1. The House-Tree-Person test requiring the teacher to draw these three objects. Scored by an experienced clinician for: Achievement Need, Control Need, Affiliation Need, Recognition Need and Ego Strength--hereafter HTP.

2. A Projective Measure of Teacher Behavior requiring the teacher to fill in the likely teacher comments to student reactions as portrayed in a series of cartoons and scored for Control and Affiliation Needs--hereafter referred to as PSI.

3. The Teacher Preference Schedule shown in previous work to be useful--hereafter TPS.

4. A Questionnaire Measure related to attitudes, philosophy, etc.--hereafter 'Objectives.'

5. The reactions to situations measure previously described*--hereafter 'Situations.'

6. The Edwards Personal Preference Schedule--hereafter EPPS.

All devices were included in a packet given individually to the teachers. Each of the tests was gone over in detail and the order of 'test taking' spelled out. The order was to be as indicated above. The teachers were permitted to fill out the instruments at their own convenience. Upon return the tests were examined as to completeness and correctness of procedure. In only a very few cases was the delay in return greater than 2 weeks. The only complaints registered by the teachers related to items on the EPPS pertaining to sex (some of the teachers were elderly) and the amount of time expended by a few of them.

---

*See Travers, R. M. W., et. al., 1961.

In addition to test data, the following biographical data was obtained on most of the teachers from district files: Age, years in teaching, college attended (since most were graduates of Utah colleges), college quarter hours past bachelors degree, year bachelors degree awarded, highest academic degree attained, and undergraduate major.

Data Processing

The first task with regard to data processing was the scoring of all materials. In most instances the pupil data required hand scoring since IBM answer sheets were used only for the end-of-year testing in the third grade. The scoring of achievement tests, questionnaires, the Barron-Welsh and sociometrics was done objectively according to keys at hand. The Circles Test is not completely objective in its scoring and hence required the training of two scorers. After training, the scorers were able to agree very well (correlations above .90). Following the scoring of all devices, data were punched on IBM cards and ready for statistical analysis.

Statistical Analysis

Statistical analysis took two forms. The first form consisted of obtaining a correlation matrix for each of the two grades and incorporating all of the measures obtained. Thus, each matrix contained each of the measures pertaining to the teacher, both of a test nature and an observational nature and also included the mean score of the class on the pupil measures. Thus, the correlations permitted the study of questions as to the predictability of behavior in the classroom and the effects of both teacher classroom behavior and teacher test scores upon change in class mean score on the various measures. In the latter instance the technique used was to examine the correlation of teacher measures with each of the post-test scores on the pupil class variables and where correlations of magnitude emerged to then partial out the pre-test score for the class by means of partial correlation. In order to reduce the total number of variables to 150 (the maximum allowable on the computer program) certain restrictions on the data were imposed. These consisted first in taking only the total ratings assigned by the raters on the Q-sort measures. Although each observer had independently Q-sorted the teacher, for this analysis the combined and joint measures were used. Similarly for the ratings made by each observer, the total rating given by two observers was used but not the individual rating of each. Finally, only the total score was utilized on Torrance Circles Test. A listing of the variables utilized with each class may be found in Appendix B. In addition to obtaining Pearson correlations among these variables scatter plots were constructed for those relationships of particular interest.

The second basic form of analysis consisted of analysis of variance applied to combinations of teacher characteristic and pupil characteristic pre-test as main effects and utilizing a variety of pupil measures as the dependent variable or score. Such analyses are not as efficient as correlational analyses for testing the relationship between the main effect and the dependent variable but do permit study of the very important interaction possibilities. Of particular interest was the relationship between teacher 'controlling' behavior as assessed through observations and two indexes of pupil characteristic at the beginning of the year; namely, achievement or ability level and dependency. To this end observation categories six and seven, (academic control and personal) were combined and this total score utilized to select approximately the highest and lowest third of teachers in each grade with respect to this dimension. Thus, the ten highest and ten lowest teachers in each of the grades were selected. For purposes of cross-validation of the analysis, these groups were sub-divided into groups of five each. Thus, for each analysis, five of the highest controlling teachers were compared with five of the lowest controlling teachers. Within each of these classes students were selected (two separate analyses) first on the basis of their fall vocabulary scores and secondly on the basis of their fall dependency questionnaire scores. It was thought important to keep analyses by sex separate; hence, this was done. Thus, within each of the classes the highest and lowest scoring boys were selected for one analysis and the highest and lowest scoring girls for separate analysis. Once the selections had been made it was possible to obtain analysis of variance on a number of dependent variables of interest, in particular the variables assessing extent of gain on several measures.

The necessity of assessing gain required that individual gain scores be obtained for pupils on the measures of interest. Accordingly a prerequisite to the analysis of variance was the obtaining of gain scores for each of the pupils. This was done by obtaining a correlation matrix of all pupil measures within each grade, setting up a first order regression equation for predicting post-test from pre-test; using this equation to predict the post-test score for each student and then subtracting this from his actual post-test score to obtain the regressed gain score. This score was then punched into the cards for each student. The pre-post- correlations and regression equations used are presented in Appendix C which also contains the correlations which we had obtained in our previous study for the same grades. It will be noted that the correlations obtained in our present study are very similar to those obtained in the previous study in grade three. In grade one, however, there are two rather striking differences. The correlation of the spring vocabulary test with the pre-test or fall vocabulary is very similar to that obtained in the past. With regard to both reading comprehension and arithmetic scores in the spring, however, the prediction based on the fall vocabulary is much poorer than we had found before. This decrease in prediction does not seem to be attributable to decrease in variability on any of the measures and hence remains an unexplained result. One can question whether or not it is necessary to adjust for pre-test scores correlating in the 20's and 30's. In order to simplify computer analyses, gain scores for only those variables indicated were punched.

CHAPTER III

MEASURES OF TEACHER CLASSROOM BEHAVIOR

Analysis of 'Observations' Data

Of paramount importance in a study relying heavily on data collected by observers in the classroom is the question of how well observers agree. Our data on the 'observations' consists of three types. First we have data pertaining to situations in which several observers observed the same teacher at the same time. Data are available for situations shortly before the onset of the major data collection and approximately half-way through. Due to the small number of observations involved, however, these data are purely descriptive. Profiles for two teachers (non-participants in the study proper) are shown in Figures 3.01 through 3.03. These data were obtained mid-way through the collection of observations and show adequate agreement among observers as well as differences in teacher behavior in the two observation sessions.

The second type of data consists of correlations between the two members of each observation team on the one occasion where they visited each teacher jointly. These correlations are presented in Table 3.01, and, again, indicate good correspondence on the major categories between different observers viewing the same teacher behavior. Two categories, 'hostility and reprimands' and 'ignores child' occur infrequently and hence exhibit too little variability in one observation period to permit assessment of observer agreement. Categories 2a--'non-verbal affiliation'--and 7m--'moralizing'--are also subject to this limitation though not so severely.

FIGURE 3.01

AGREEMENT OF FOUR OBSERVERS--TEACHER #1--FIRST GRADE

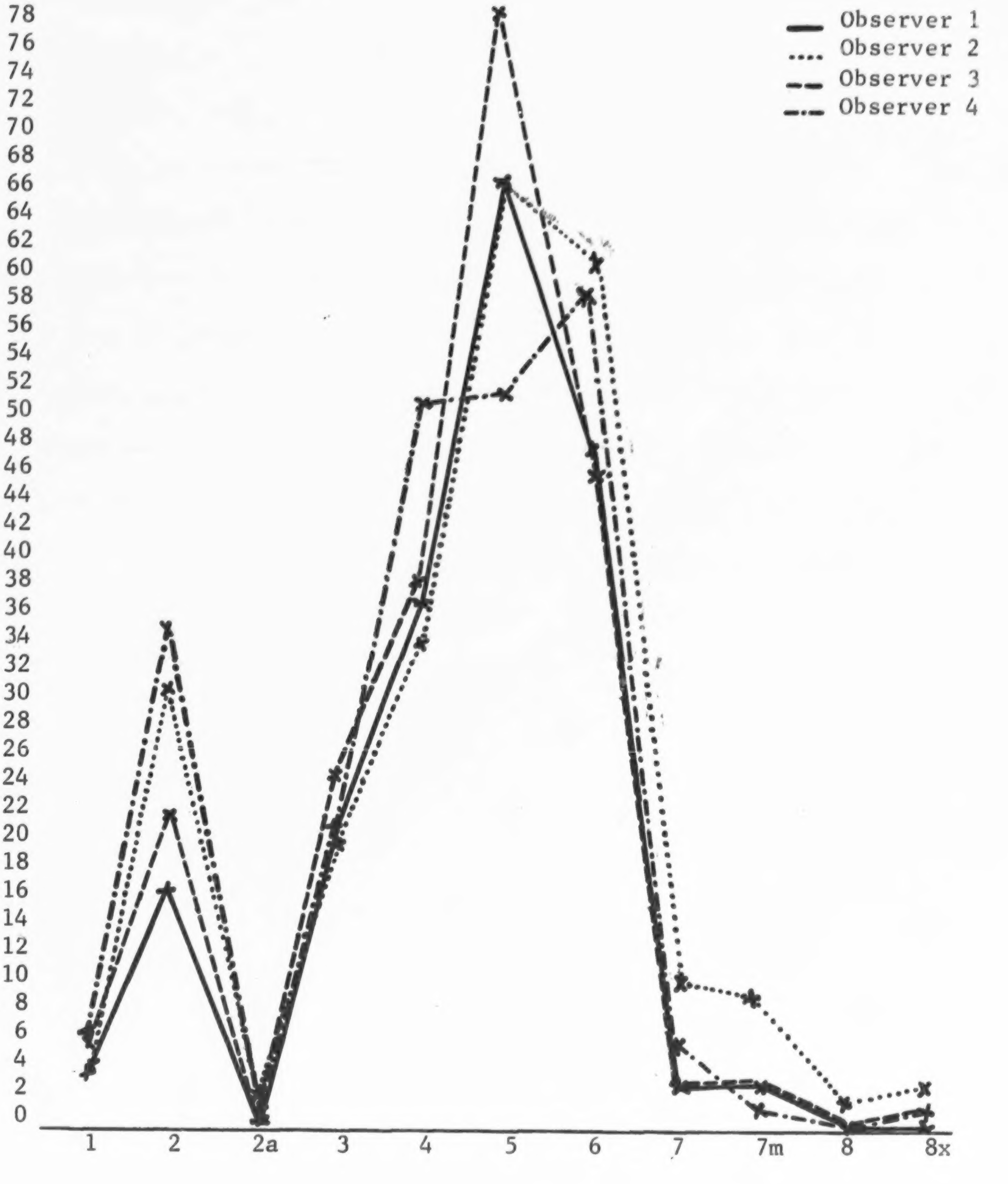


FIGURE 3.02

AGREEMENT OF FOUR OBSERVERS--TEACHER #2--SECOND GRADE

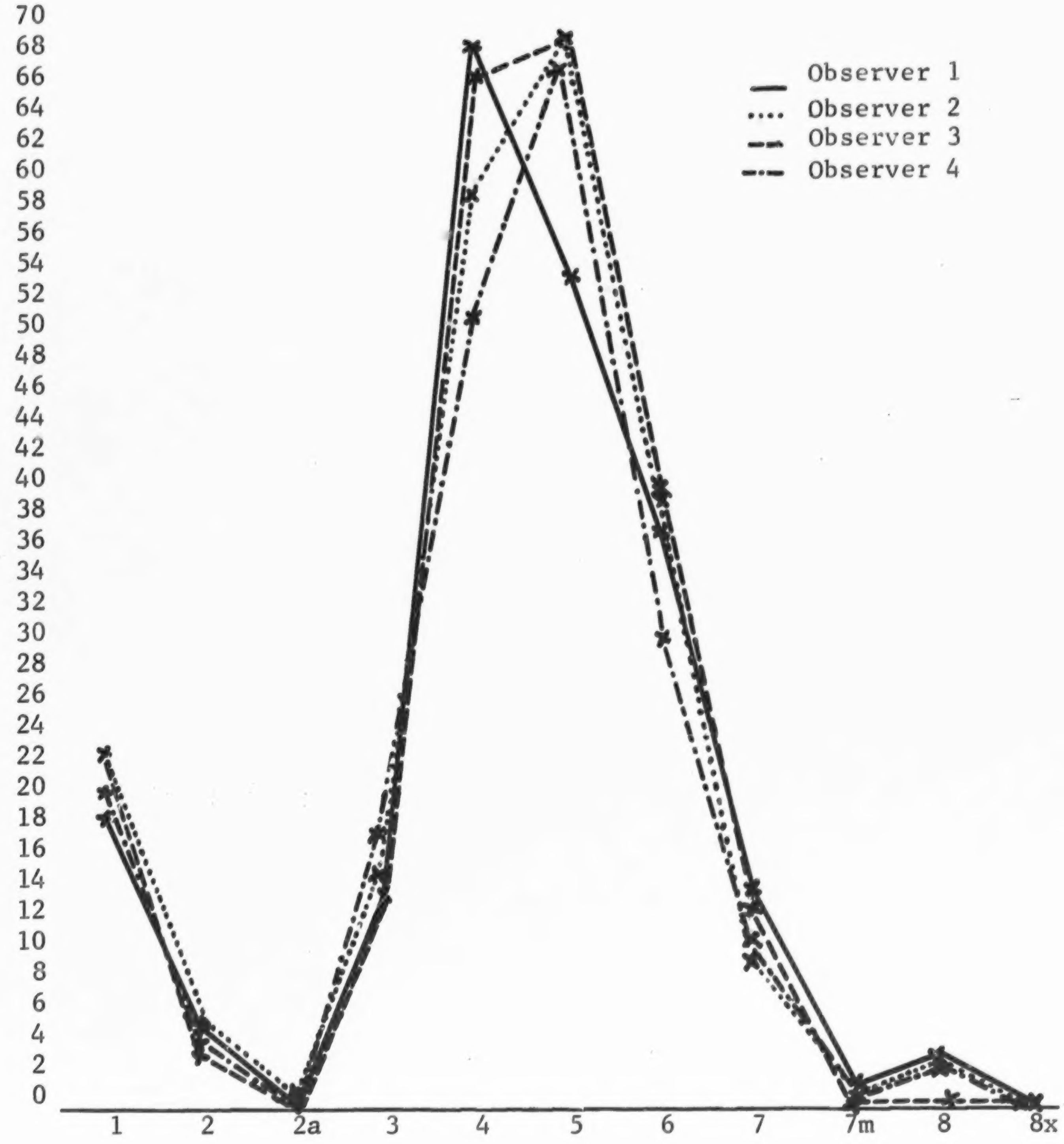


FIGURE 3.03

COMPARISON OF TWO TEACHERS ON 'OBSERVATIONS'--SCORES ARE MEANS BASED ON FOUR OBSERVERS

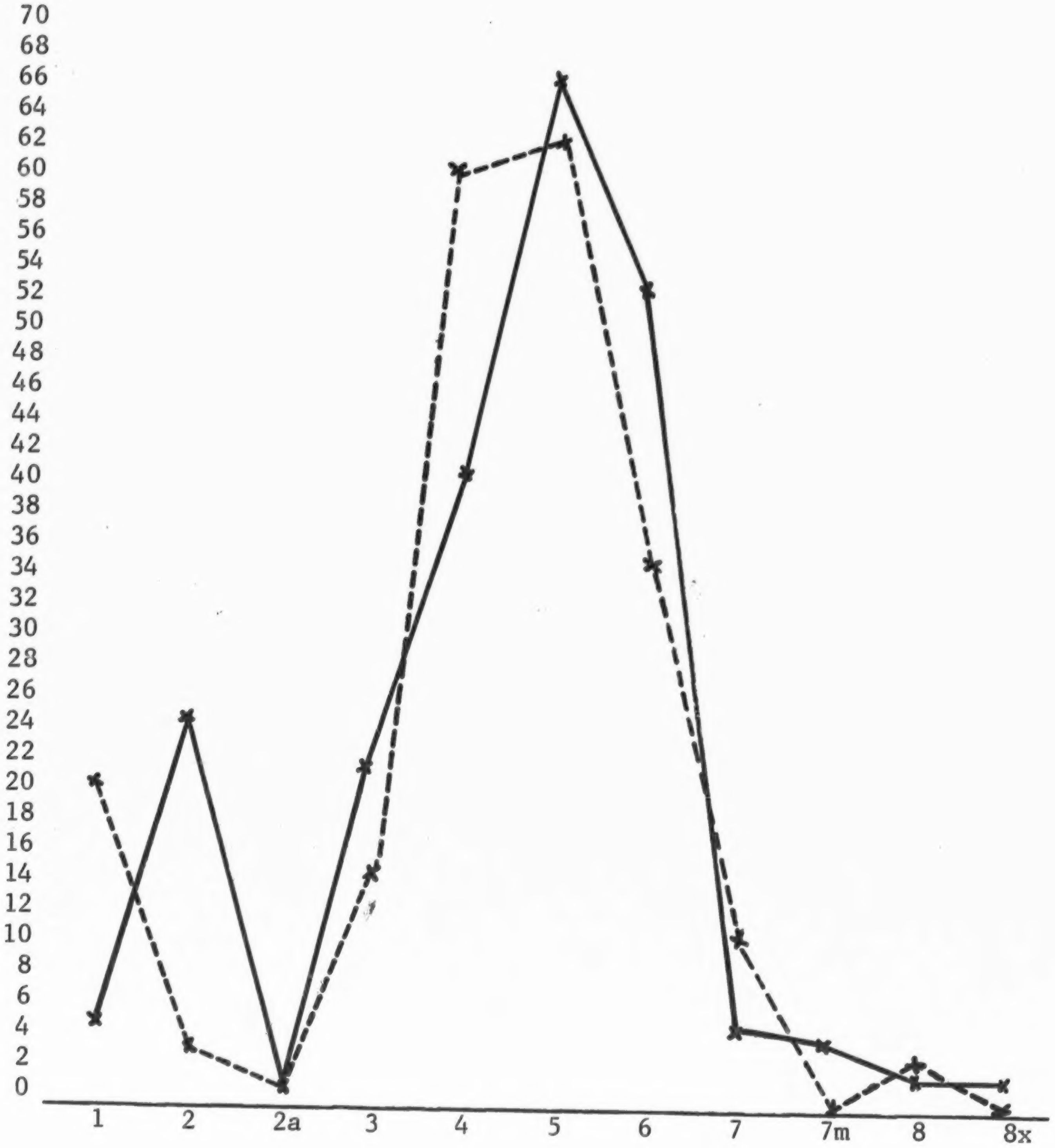


TABLE 3.01

AGREEMENT AMONG THE TWO MEMBERS OF EACH OBSERVATION TEAM ON INDEPENDENT SCORING OF SAME TEACHING SITUATION

Category	Team I N = 33 Teachers	Team II N = 34 Teachers
1. Acknowledges Hand	.96	.93
2. Praise and Encouragement	.89	.92
2a. Non-verbal Affiliation	.95	.85
3. Minimal Reinforcement	.86	.82
4. Asking Questions	.96	.94
5. Problem Structuring	.88	.76
6. Academic Control	.94	.86
7. Personal Control	.96	.92
7m. Moralizing	.85	.65
8. Hostility and Reprimands	---	---
8x. Ignores Child	---	---

The third type of data available consists of analysis of observations made by the observers for the teacher samples. Analysis of variance was performed for each of the major observation categories, i.e., categories two through eight. The analysis for each variable was performed separately for each of the two teams of teachers since each team had visited a distinct group of teachers. These analyses are shown in Tables 3.02 through 3.08. It will be noted that these analyses do not include interaction terms; that is, the residual term could be further subdivided into the interaction between teachers and observers which could be further

TABLE 3.02

ANOVA--OBSERVATIONS--CATEGORY 2, 'PRAISE AND ENCOURAGEMENT'

Source	df	Team I SS	MS	F	p
Between Teachers (35)		(8088.84)	(231.11)	6.42	
Between Grades	1	15.87	15.87	<1.00	N.S.
Between Schools	8	2532.24	316.53	1.48	N.S.
Residual b/t Teachers	26	5540.73	213.10	5.87	$<.001$
Between Observations (7)		(676.97)			
Between Observers	1	348.92	348.92	9.61	$<.01$
Between Sessions	3	313.84	104.61	2.88	$<.05$
Residual: Observations	3	14.21	4.74	4.00	N.S.
Residual (245)		8894.91	36.31		
TOTAL (287)		17660.72			
Teacher Variance % = 47%					
		Team II			
Between Teachers (39)		(4554.79)	116.79	3.21	
Between Grades	1	133.92	133.92	<1.00	N.S.
Between Schools	8	371.31	46.41	<1.00	N.S.
Residual b/t Teachers	30	4049.56	134.99	3.72	$<.001$
Between Observations (7)		(1461.55)			
Between Observers	1	322.00	322.00	8.86	$<.01$
Between Sessions	3	971.06	323.69	8.91	$<.001$
Residual Observations	3	168.49	56.16	1.54	N.S.
Residual (273)		9919.38	36.33		
TOTAL (319)		14474.17			
Teacher Variance % = 32%					

TABLE 3.03

ANOVA--OBSERVATIONS--CATEGORY 3, 'MINIMAL REINFORCEMENT'

		Team I			
Source	df	SS	MS	F	p
Between Teachers (35)		(7090.09)	(202.57)	3.71	
Between Grades	1	803.04	803.04	7.91	<.01
Between Schools	8	3647.47	455.93	4.49	<.001
Residual b/t Teachers	26	2639.58	101.52	1.86	<.01
Between Observations (7)					
Between Observers	1	592.25	592.25	10.84	<.001
Between Sessions	3	125.68	41.89	<1.00	N.S.
Residual: Observations	3	85.01	28.34	<1.00	N.S.
Residual (245)		13385.44	54.63		
TOTAL (287)		21278.47			
Teacher Variance % = 32%					
		Team II			
Between Teachers (39)		(4233.74)	(108.56)	2.30	
Between Grades	1	2553.58	2553.58	61.93	<.001
Between Schools	8	1556.47	194.56	4.72	<.001
Residual: Teachers	30	123.69	41.23	<1.00	N.S.
Between Observations (7)		(502.59)			
Between Observers	1	110.45	110.45	2.34	N.S.
Between Sessions	3	111.44	37.15	<1.00	N.S
Residual: Observations	3	280.66	93.55	1.98	N.S.
Residual (273)		12911.66	47.30		
TOTAL (319)		17647.99			
Teacher Variance % = 11%					

TABLE 3.04

ANOVA--OBSERVATIONS--CATEGORY 4, 'ASKING QUESTIONS'

Source	df	Team I SS	MS	F	p
Between Teachers (35)		(37877.53)	(1082.22)	6.42	
Between Grades	1	4453.09	4453.09	5.78	<.05
Between Schools	8	13395.74	1674.47	2.17	<.05
Residual b/t Teachers	26	20028.70	770.33	4.57	<.001
Between Observations (7)		(2885.66)			
Between Observers	1	547.25	547.25	3.25	N.S.
Between Sessions	3	686.70	228.90	1.36	N.S.
Residual: Observations	3	1651.71	550.57	3.27	<.05
Residual (245)		41279.72	168.49		
TOTAL (287)		82042.91			
Teacher Variance % = 43%					
		Team II			
Between Teachers (39)		(49601.69)	(1271.84)	4.10	
Between Grades	1	289.26	289.26	<1.00	N.S.
Between Schools	8	14443.08	1805.38	1.55	N.S.
Residual: Teachers	30	34869.35	1162.31	3.74	<.001
Between Observations (7)		(3601.19)			
Between Observers	1	1540.01	1540.01	4.96	<.05
Between Sessions	3	376.11	125.37	<1.00	N.S.
Residual: Observations	3	1685.07	561.69	1.81	N.S.
Residual (273)		84783.81	310.56		
TOTAL (319)		137986.69			
Teacher Variance % = 23%					

TABLE 3.05

ANOVA--OBSERVATIONS--CATEGORY 5, 'PROBLEM STRUCTURING'

Source	df	Team I SS	MS	F	p
Between Teachers (35)		(15459.88)	(441.71)	3.20	
Between Grades	1	960.21	960.21	3.05	N.S.
Between Schools	8	6304.46	788.06	2.50	$<.05$
Residual Teachers	26	8195.21	315.20	2.28	$<.001$
Between Observations (7)		(9734.97)			
Between Observers	1	8011.67	8011.67	57.97	$<.001$
Between Sessions	3	1138.96	379.65	2.75	$<.05$
Residual: Observations	3	584.34	194.78	1.41	N.S.
Residual (245)		33856.65	138.19		
TOTAL (287)		59051.50			
Teacher Variance % = 29%					
		Team II			
Between Teachers (39)		(57278.45)	(1468.67)	4.91	
Between Grades	1	413.26	413.26	<1.00	N.S.
Between Schools	8	19389.29	2423.66	1.94	N.S.
Residual: Teachers	30	37475.90	1249.20	4.18	$<.001$
Between Observations (7)		(2692.10)			
Between Observers	1	812.81	812.81	2.72	N.S.
Between Sessions	3	1468.18	489.39	1.64	N.S.
Residual: Observations	3	411.11	137.04	<1.00	N.S.
Residual (273)		81648.65	299.08		
TOTAL (319)		141619.20			
Teacher Variance % = 41%					

TABLE 3.06

ANOVA--OBSERVATIONS--CATEGORY 6, 'ACADEMIC CONTROL'

		Team I			
Source	df	SS	MS	F	p
Between Teachers (35)		(88.07)	(251.63)	2.26	
Between Grades	1	1217.91	1217.91	1.56	<.05
Between Schools	8	2459.90	307.49	1.56	N.S.
Residual b/t Teachers	26	5129.19	197.27	1.77	<.05
Between Observations (7)		(6498.22)			
Between Observers	1	4851.12	4851.12	43.60	<.001
Between Sessions	3	390.08	130.03	1.17	N.S.
Residual: Observations	3	1257.02	419.01	3.77	<.05
Residual (245)		27260.78	111.27		
TOTAL (287)		42566.00			
Teacher Variance % = 21%					
		Team II			
Between Teachers (39)		(26213.97)	(672.15)	3.01	
Between Grades	1	735.86	735.86	1.55	N.S.
Between Schools	8	11248.11	1406.01	2.96	<.01
Residual: Teachers	30	14230.00	474.33	2.12	<.001
Between Observations (7)		(915.13)			
Between Observers	1	564.46	564.46	2.53	N.S.
Between Sessions	3	192.24	64.08	<1.00	N.S.
Residual: Observations	3	158.43	52.81	<1.00	N.S.
Residual (273)		61014.75	223.50		
TOTAL (319)		88143.85			
Teacher Variance % = 29%					

TABLE 3.07

ANOVA--OBSERVATIONS--CATEGORY 7, 'PERSONAL CONTROL'

		Team I			
Source	df	SS	MS	F	p
Between Teachers (35)		(8578.82)	(245.11)	6.72	
Between Grades	1	477.31	477.31	1.92	N.S.
Between Schools	8	1638.90	204.86	<1.00	N.S.
Residual b/t Teachers	26	6462.61	248.56	6.82	<.001
Between Observations (7)		(1411.49)			
Between Observers	1	1120.22	1120.22	30.73	<.001
Between Sessions	3	87.24	29.08	<1.00	N.S.
Residual: Observations	3	204.03	68.01	1.87	N.S.
Residual (245)		8930.50	36.45		
TOTAL (287)		17509.32			
Teacher Variance % = 51%					
		Team II			
Between Teachers (39)		(16235.49)	(416.29)	12.25	
Between Grades	1	626.25	626.25	1.98	N.S.
Between Schools	8	6133.48	766.68	2.43	<.05
Residual: Teachers	30	9475.76	315.86	9.30	<.001
Between Observations (7)		(183.69)			
Between Observers	1	42.05	42.05	1.23	N.S.
Between Sessions	3	18.66	6.22	<1.00	N.S.
Residual: Observations	3	122.98	40.99	1.21	N.S.
Residual (273)		9277.81	33.98		
TOTAL (319)		25696.99			
Teacher Variance % = 62%					

TABLE 3.08

ANOVA--OBSERVATIONS--CATEGORY 8, 'HOSTILITY AND REPRIMANDS

		Team I			
Source	df	SS	MS	F	p
Between Teachers (35)		(595.70)	(17.02)	3.79	
Between Grades	1	27.22	27.22	1.59	N.S.
Between Schools	8	122.37	15.30	<1.00	N.S.
Residual b/t Teachers	26	445.61	17.14	3.82	<.001
Between Observations (7)		(15.05)			
Between Observers	1	2.53	2.53	<1.00	N.S.
Between Sessions	3	3.65	1.22	<1.00	N.S.
Residual: Observations	3	8.87	2.96	<1.00	N.S.
Residual (245)		1099.33	4.49		
TOTAL (287)		1710.08			
Teacher Variance % = 34%					
		Team II			
Between Teachers (39)		(1526.88)	(39.15)	8.28	
Between Grades	1	8.66	8.66	<1.00	N.S.
Between Schools	8	233.81	29.23	<1.00	N.S.
Residual: Teachers	30	1284.41	42.81	9.05	<.001
Between Observations (7)		(59.28)			
Between Observations	1	48.83	48.83	10.32	<.001
Between Sessions	3	6.84	2.28	<1.00	N.S.
Residual: Observations	3	3.61	1.20	<1.00	N.S.
Residual (273)		1193.39	4.73		
TOTAL (319)		2771.50			
Teacher Variance % = 56%					

TABLE 3.09

COMPARISON OF TWO OBSERVER TEAMS ON MEANS IN MAJOR OBSERVATION CATEGORIES

	Observer A	Team I Observer B	Total
2. Praise and Encouragement	12.13	9.93	11.03
3. Minimal Reinforcement	16.86	13.99	15.43
4. Asking Questions	40.40	37.64	39.02
5. Problem Structuring	43.03	32.48	37.75
6. Academic Control	35.44	27.23	31.33
7. Personal Control	11.92	7.98	9.95
8. Hostility Reprimands	1.12	1.31	1.22

	Observer C	Team II Observer D	Total
2. Praise and Encouragement	7.11	9.11	8.11
3. Minimal Reinforcement	15.11	16.28	15.69
4. Asking Questions	42.02	46.41	44.22
5. Problem Structuring	60.62	63.79	62.20
6. Academic Control	31.15	33.81	32.48
7. Personal Control	10.58	11.31	10.94
8. Hostility Reprimands	.96	1.74	1.35

subdivided into interactions between grades and observers; between schools and observers and a residual between teachers and observers. Further, an interaction between teachers and situations, i.e., time sequencing, could be obtained which could be subdivided also into a grade by situation, a school by situation, and a residual teacher by situation. Interactions were not separated out of the residual partially because of the requirements of computing time, partly because such interactions are of little interest in themselves, and partly because it was considered that a residual term which left these components in was most meaningful.

Perhaps the most interesting result of such analyses pertains to the extent to which between teacher differences are significant and assume a major portion of the total variance. It will be noted that the residual between teacher's term which consists of the variation remaining among teachers after the between grades and between schools components have been separated out is significant and of considerable proportion in all categories with the exception of variable number three, 'minimal reinforcement.' For this variable the residual 'between teacher' component is non-significant for one team and barely significant for the other team. In this instance there is some question as to whether or not a major portion of the variance is due to differences in observed behavior of the teachers. A question, however, is whether or not the 'between schools' variation should be combined with the 'between teachers' since the between schools is in this and some other analyses significant. If one feels that the difference observed between schools is primarily a function of differences in teacher behavior as such, then one can justify its inclusion within such a term. If on the other hand one feels that this is not a typical picture of the teacher's behavior but rather is a function of school policy or the school mileau, then presumably one should partial it out. In any case, there is some question as to the usefulness of this particular observation category ('minimal reinforcement') in differentiating teachers.

The second question of importance is the extent to which the two team members agree overall in their appraisal of teachers, as reflected in their scoring. For team one it will be noted that there are a number of significant between observer differences; frequently of large magnitude. In all cases these reflect a propensity for one observer to score more frequencies in a particular category than the other. This is true in five of the seven categories. This factor should not introduce bias in subsequent analyses; however, since the total impressions of both observers are summed to provide the major score of teacher behavior and since, for the frequency score (f), raw data were converted to derived scores based on the distribution for each observer before summing. For the other team two of the categories indicate significant differences between the observers at the .01 level indicating that one observer typically records a higher frequency in two categories--'praise and encouragement' and 'hostility and reprimands'--suggesting that this observer was more sensitive to both extremes of the continuum between hostility and supportive behavior. Once again, since both of these are in the same direction, this should not distort the final scores used.

The next consideration has to do with the degree to which the two teams are observing and scoring behavior similarly. Our data do not provide conclusive evidence on this point because the two teams observed different groups of teachers and it is quite conceivable that the teacher groups may have differed in some ways though this is not felt to be likely. Table 3.09 compares the two teams with regard to overall mean frequencies in each category. It will be noted that the team scores are in close agreement in all categories except one--the problem structuring category wherein team two has a mean value which is almost twice that of team one.

Thus, for some reason it appears that this team was much more responsive to problem structuring behavior on the part of the teacher. Because of this divergency this particular category must be treated very tentatively in subsequent analyses.

Although not of primary importance to the present study, two other components of the analyses of variance are of interest. These relate to between grade differences and between sequences differences. The latter does not affect the overall impression of teacher behavior and the former is not involved in other analyses since all were conducted separately for each grade. As to between grades comparisons, the major categories in which clear differences emerged are: category three, 'minimal reinforcement,' and category seven, 'personal control.' Both teams observed much more minimal reinforcement in grade three. On category seven one team scores more in grade one and the other more in grade three. The category in which there is clearly no important between grade differences is category number two, 'praise and encouragement' in both grades.

One further use was made of these analyses. It is of considerable interest to estimate the proportion of the total variation which may be attributed to systematic differences between teachers when the individual observations were made by two observers over a total of 8* observation periods. It is obviously true that any teacher would not receive precisely the same set of frequencies on different days and a common question which arises is the extent to which differences which are observed are in fact differences between teacher's typical behavior rather than differences among observation visits. A very crude estimate of this proportion was attained as follows: Beginning with the sum of squares as reported in these tables, the proportion of the observed variance attributable to between teacher differences was estimated as follows First, it is argued that it is legitimate to combine the sum of squares of the 'between schools' and the 'residual between teachers' since in many cases the between school factor is not significant when compared with the residual between teachers and since, where it is significant, one can argue that the school differences are simply artifacts of the teachers within those schools. Secondly, the total sums of squares was reduced by two factors: the between grades sum of squares and the between observer's sum of squares. This is held to legitimate since we are interested really in the amount of between teacher differences which would exist within a given grade and also when observer differences of a systematic sort are balanced out. Thus, in effect the proportion of the total variance less observer and grades differences which is constituted by the between teacher's variance is obtained. It will be noted that the total variance then is allowed to contain all of the other residual components including some interactions with grade and observer which it could be argued should be deleted. Hence, the final proportions probably give an underestimate of the proportion of total variance which may be legitimately attributed to variance among teachers within a given grade and after observer differences are controlled.

---

*The joint observation is not included in these analyses.

It may be noted in Tables 3.02 through 3.08 that the percentages vary somewhat from one team of observers to the other. This is most likely due in part to the interaction of observer differences with other variables which is allowed to remain in the residual term but may also reflect differences in efficiency of the two teams on certain variables. The percentage of total variance attributable to 'between teacher' differences for each of the categories two through eight is roughly as follows: 2--40 percent, 3--20 percent, 4--35 percent, 5--35 percent, 6-25 percent, 7--55 percent and 8--45 percent. Crude though these figures are they suggest that for each of the observational categories a substantial proportion of the differences in each category observed on different sessions is attributable to differences among teachers; the remaining variance presumably being attributable to fluctuation in teacher behavior from day to day and possible interactions of observers with teachers with particular days of observation.

Interrelationships Among Measures of Teacher Classroom Behavior

The extent to which the various indices of teacher classroom behavior agree with one another is a matter of considerable importance. It will be recalled that three different though not independent techniques were used in the assessment of teacher behavior. The most objective of these measures was the tallying of teacher behavior in each of 11 categories during each 30-minute observation session. Totals across the nine observation sessions within each category provided the basis of first measure of teacher behavior. A modification of this procedure resulted in a second score which simply translated each of these total frequencies into a percentage based on the total number of tallies recorded for each teacher since teachers differed in the amount of categorizable behavior observed. The second measurement approach was a summation of the rating assigned by the observer immediately after each observation period in each of seven categories. The third measure consisted of Q-sorts completed after all of the observations had been completed. Two variations of this were utilized. The first score is a total score which is the sum of the Q-sort scores assigned independently by the two observers and the second is a score assigned by the two observers Q-sorting together.

Agreement Between Independent and Joint Q-sorts

Table 3.10 shows the correlation between the Q-sorts under the two conditions. Ideally, of course, one would hope for correlations near 1.0 but the resultant correlations show considerable agreement between the two approaches to Q-sorting. For the first grade, all of the correlations are above .75 and most are above .80. For grade three the correlations are, for most variables, very similar to those found in grade one, the major exception being that the stimulating dimension is much less consistent; whereas the correlation was .84 in grade one it is only .55 in grade three placing a rather severe limitation upon this particular measure for the third grade.

Relationships Between Q-Sorts (Joint) and Ratings

The next question concerns the degree of agreement between the Q-sorts and the rating procedures. For purposes of simplicity only the 'joint' Q-sort correlations are presented. The correlations based on 'total' Q-sort are very similar. Once again the extent of agreement as shown in Tables 3.11 and 3.12 is gratifying. The correlations between similar or identical variables under the two procedures are in general quite high; whereas correlations with the other variables tend to be considerably lower. Thus, for example, in grade one the correlation between the two measures of supportive behavior is .85, between the two measures of punitive behavior .81, between the measures of affiliation or warmth in behavior .76, and between the two measures of stimulating behavior .71. The correlation with controlling behavior is somewhat lower than this at .65.

TABLE 3.10

CORRELATIONS BETWEEN 'JOINT' AND 'TOTAL' Q-SORTS GRADES I AND III

	Grade I	Grade III
Controlling	.75	.84
Warm	.78	.81
Punitive	.92	.88
Confident	.76	.75
Supportive	.83	.85
Stimulating	.84	.55
Achievement Oriented	.80	.81
Physical Control	.78	.70
Intellectually Effective	.86	.82
Disparaging	.81	.89
Factor I	.93	.89

It is pertinent here to point out that several of the measures both in the Q-sorts and ratings would be expected on logical grounds to correlate fairly highly; that is, one would expect correlations between the warmth of the teacher and the extent of supportive behavior, and between the extent of punitive behavior and disparaging behavior. These correlations do in fact exist both within the ratings and the Q-sorts and in general range above .75 and in a coherent direction.

In grade three as was the case with grade one, quite good agreement exists between Q-sort (joint) and ratings and in all cases the correlations which are expected are higher than those with variables which

TABLE 3.11

CORRELATIONS BETWEEN Q-SORTS (JOINT) AND RATINGS--GRADE I

	RATINGS						
Q-SORTS	Controlling	Stimulating	Disparaging	Supportive	Confident	Affiliative	Relaxed
Controlling	.65	-.25	.68	-.55	-.26	-.42	.06
Warm	-.30	.48	-.59	.77	.28	.76	-.23
Punitive	.49	-.51	.81	-.86	-.36	-.75	-.19
Confident	.02	.18	-.06	-.03	.72	-.05	-.14
Supportive	-.34	.62	-.66	.85	.30	.85	-.42
Stimulating	-.11	.71	-.39	.44	.51	.48	-.70
Achievement Oriented	.20	.36	.08	-.13	.34	-.14	-.57
Physical Contact	-.07	.13	-.09	.29	.06	.37	.20
Intellectually Effective	.03	.62	-.22	.16	.30	.14	-.70
Disparaging	.45	-.49	.68	-.69	-.46	-.52	.18
Factor I	.52	-.54	.80	-.89	-.35	-.81	.26

TABLE 3.12

CORRELATIONS BETWEEN Q-SORTS (JOINT) AND RATINGS--GRADE III

	RATINGS						
Q-SORTS	Controlling	Stimulating	Disparaging	Supportive	Confident	Affiliative	Relaxed
Controlling	.60	-.24	.66	-.44	-.08	-.50	-.19
Warm	-.44	.28	-.56	.64	.11	.76	.15
Punitive	.70	-.40	.82	-.56	-.06	-.56	-.05
Confident	.23	.38	.11	-.03	.52	-.10	-.48
Supportive	-.53	.36	-.67	.68	.18	.74	.06
Stimulating	-.35	.61	-.39	.41	.52	.42	-.41
Achievement Oriented	.14	.37	.12	-.05	.30	-.24	-.59
Physical Contact	.09	.05	-.15	.36	.06	.49	.30
Intellectually Effective	-.25	.53	-.26	.20	.47	.06	-.54
Disparaging	.62	-.38	.81	-.50	-.07	-.53	-.04
Factor I	.65	-.35	.76	-.61	-.11	-.67	-.15

are not intended to be the same or similar. The level of correspondence is in general somewhat lower than was true in grade one. As was true in grade one, the best agreement between ratings and Q-sorts occurs on the punitive and supportive and warmth scales with somewhat lower correlations for the controlling and stimulating dimensions. The agreement between Q-sorts and ratings on the confidence dimension is considerably lower (.52) than was the case in grade one where the correlation was .72. The correlations between the Q-sort of 'intellectually effective' and the ratings stimulating and relaxed are somewhat lower than was the case in grade one, being of the magnitude of .55 rather than a magnitude of .65.

Relationships Between the Observations and Q-Sorts

These relationships are of particular interest because the data based on observational tallies are considered to be our most objective form of data and it is of considerable interest to see how well these agree with the impressions of observers as reflected in Q-sorts. Again one would hope for very high agreement on similar variables though one could not expect it to be perfect, since the observer impressions will be influenced by their other contacts with the teachers and by their reactions to teacher behavior which may not be reflected in the categories. The correlations for grade one are shown in Table 3.13. Once again the relationships are quite satisfactory. Taking each of the observational categories in turn we note that the category 'answering child's hand' is related only ($\approx$.30) to the Q-sort variable of 'confidence.' This appears logical although it is difficult to predict which of the Q-sort variables might be expected to relate to this category since it is a behavior engaged in to a large extent by all teachers and yet does not seem to have a particular psychological meaning. The second category, 'praise and encouragement,'

TABLE 3.13

CORRELATIONS BETWEEN OBSERVATIONS (f and %) AND Q-SORTS (JOINT)--GRADE I

Q-Sorts		Cont.	Warm	Pun.	Conf.	Supp.	Stim.	Ach. Or.	Phys. Cont.	Int. Effect.	Disp.
Observations:											
Answers Hand	f 84	.07	-.02	-.05	.30	.06	.07	.09	.07	.02	.12
	% 95	.02	.01	-.09	.31	.13	.09	.06	.01	.04	.10
Praise & Encouragement	f 85	-.04	.67	-.47	.01	.60	.39	-.18	.44	-.04	-.33
	% 96	-.13	.58	-.48	-.01	.51	.29	-.26	.37	-.17	-.42
Non-Verbal Affiliation	f 86	.06	.31	-.12	-.36	.26	.02	-.13	.60	-.11	-.02
	% 97	.00	.33	-.19	-.28	.30	.01	-.14	.42	-.09	.04
Minimal Reinforcement	f 87	-.12	.30	-.35	.38	.28	.28	.25	.22	.24	-.43
	% 98	-.22	.26	-.42	.48	.28	.19	.23	.14	.18	-.54
Asking Questions	f 88	-.08	.15	-.22	.07	.18	.31	.35	-.02	.51	-.15
	% 99	-.24	.01	-.26	.12	.16	.32	.45	-.01	.56	-.25
Problem Structuring	f 89	-.44	.21	-.32	-.05	.21	.14	-.15	-.01	.04	-.28
	% 100	-.44	.14	-.29	-.02	.17	.04	-.19	-.23	-.01	-.22
Academic Control	f 90	.37	-.17	.34	-.16	-.19	-.01	.16	.08	.02	.18
	% 101	.21	-.31	.36	-.15	-.30	-.15	.10	.04	-.10	.15
Personal Control	f 91	.64	-.57	.78	-.21	-.62	-.44	-.11	.01	-.35	.70
	% 102	.59	-.43	.74	-.22	-.59	-.44	-.19	-.08	-.35	.66
Moralizing	f 92	.36	.05	.27	-.06	-.11	-.18	-.03	.34	-.13	.31
	% 103	.24	.14	.13	-.17	.02	-.15	-.15	.12	-.10	.18
Hostility & Reprimands	f 93	.67	-.59	.80	-.12	-.67	-.44	-.08	-.29	-.23	.71
	% 104	.60	-.41	.70	-.13	-.60	-.40	-.13	-.22	-.27	.64
Ignores Child	f 94	.33	-.08	.33	-.20	-.27	-.28	-.19	-.11	-.12	.32
	% 105	.29	-.12	.34	-.13	-.33	-.32	-.19	-.14	-.14	.34

relates positively, as would be expected, to both the Q-sort on 'warmth' and the Q-sort on 'supportive' behavior; the correlations being of the magnitude of .60 and also relates to a somewhat lesser degree (-.47) to the Q-sort on punitive behavior and again to a slightly lesser degree to the Q-sort on physical contact (≈.40). The next category, non-verbal affiliation, relates primarily to the Q-sort on physical contact, the correlation being of the magnitude of .40 to .60 and to a lesser degree (magnitude of about .30) with warmth and negatively with confidence. The next category, 'minimal reinforcement,' does not show particularly high correlations with any of the Q-sorts,the highest being approximately .40 with 'confidence' and in the 30's negative with punitive behavior. The category, 'asking questions,' is not highly related to any of the Q-sorts, the highest correlation with the magnitude of .30 being with 'stimulating' and 'achievement oriented' which makes sense. The category, 'problem structuring,' is correlated in the magnitude of .40 negatively with controlling Q-sort. The category, 'academic control,' shows correlations of the magnitude of about .30 with controlling Q-sort which is somewhat lower than might have been expected, and correlations of the magnitude of .30 with punitiveness. The latter is not particularly to be expected but suggests that the teacher who exerts the more academic control is viewed by the observers as being somewhat more punitive. The next category, 'personal control,' shows very substantial correlations with several of the Q-sorts. Correlations in the magnitude of .60 with controlling Q-sort suggesting that the Q-sort control variable is primarily a function of the personal control exerted by the teacher rather than academic control. Correlations of the magnitude of .75 are found with the 'punitive' category and negative correlations of the magnitude of -.60 with the 'supportive'

Q-sort, and of the magnitude of -.50 with the 'warmth' Q-sort. Thus, it appears that the dimension of personal control as exerted by the teacher is a very important dimension in affecting the Q-sorts, at least in the first grade, in that the teacher who exerts a great deal of personal control over the children's behavior is viewed as controlling, punitive, cold, non-supportive, and to a lesser extent (.40) non-stimulating. The category, 'moralizing,' is one having a relatively low frequency of occurrence and consequently relatively low variability across teachers and shows only modest correlations of the magnitude of .30 with controlling Q-sort. The category, 'hostility and reprimands,' shows several rather high correlations. Correlations of .75 magnitude are found with the punitive Q-sort and correlations of the magnitude of .60 with the controlling Q-sort with non-supportive behavior and with neuroticism. Finally, the category, 'ignoring child,' shows modest correlations of .30 with controlling punitive Q-sort non-supportive and non-stimulating Q-sort.

Without going into detail it is clear that one can approach this matrix from the other standpoint and look at each of the Q-sort variables and its relationship with the categories. Once again, this inspection supports the notion that both approaches, the observation and the Q-sorts, are measuring essentially the same variables which is to be desired. Beyond this, this analysis in total suggests that controlling behavior as viewed by observers is largely a function of personal control and hostility more than academic control.

It is worth making special note, however, of those variables on the Q-sorts which seem not to have direct counterparts in the observation categories. In particular it should be noted that teachers viewed as 'confident' by observers tended to engage in less non-verbal affiliation, more minimal reinforcement, and somewhat more acknowledging of children's hands, in total suggesting a responsive but not physically overt reaction toward the children. Teachers viewed as 'stimulating' by the observers tended to be low on 'personal control' and 'hostility' and tended to be higher on 'asking questions' and 'praise and encouragement,' though none of these correlations is above .45. The teacher viewed by observers as 'achievement oriented' tends to ask more questions and to give somewhat more minimal reinforcement but once again the correlations are not large.

In general the third grade correlations as shown in Table 3.14 are quite similar to the first grade correlations and in accord with expectations. Thus, the praise and encouragement observations correlate positively with the Q-sort on warmth and supportive behavior; correlate negatively with the punitive scale; and correlate .41 with the stimulating dimension. Unlike the first grade, however, the correlations with physical contact are very low. The 'non-verbal affiliation' score correlates with the physical contact Q-sort as was true in the first grade and with the warmth dimension as was also the case in the first grade. The 'minimal reinforcement' category correlates negatively with the punitive Q-sort at a level comparable to the first grade (approximately .40). The correlation with the disparaging Q-sort is approximately -.30 whereas in the first it was at the magnitude of -.50. The 'asking questions' variable shows considerably lower correlations in the third grade and in particular the correlations with the Q-sort of 'intellectual effectiveness' and with 'achievement orientation' drop considerably from a magnitude of .50 to a magnitude of .20. The correlations with the 'problem structuring' category are also lower in the third and in fact virtually none of the correlations are significant. Third grade correlations for the academic

TABLE 3.14

CORRELATIONS BETWEEN OBSERVATIONS (f AND %) AND Q-SORTS (JOINT)--GRADE III

Q-Sorts		Cont.	Warm	Pun.	Conf.	Supp.	Stim.	Ach. Or.	Phys. Cont.	Neur.	Int. Effect.	Disp.
Observations:												
Answers Hand	f 94	.06	.04	-.18	-.01	.05	.02	.14	-.28	-.02	.13	-.03
	% 105	-.02	.03	-.25	-.08	.05	.02	.12	-.30	-.02	.12	-.13
Praise & Encouragement	f 95	-.25	.47	-.35	.22	.42	.41	.02	.12	-.29	.28	-.19
	% 106	-.16	.39	-.27	-.01	.35	.30	-.02	.07	-.16	.24	-.16
Non-Verbal Affiliation	f 96	-.02	.32	.06	-.01	.26	.26	.11	.50	.19	-.09	-.06
	% 107	-.05	.38	.01	-.08	.28	.27	.03	.43	.11	-.08	-.20
Minimal Reinforcement	f 97	-.15	.17	-.47	.15	.17	.17	.02	-.18	-.57	.24	-.29
	% 108	-.16	.28	-.43	-.07	.23	.09	.02	-.18	-.51	.21	-.28
Asking Questions	f 98	.04	.11	-.04	.19	.05	.17	.01	-.04	-.18	.19	.03
	% 109	-.08	.21	-.11	.08	.18	.16	-.07	.07	-.19	.21	-.10
Problem Structuring	f 99	-.14	-.16	-.12	.31	-.01	.20	.08	-.05	-.20	.09	-.09
	% 110	-.27	-.17	-.14	.16	.02	.09	.07	.06	-.06	-.01	-.24
Academic Control	f 100	.39	-.30	.44	.08	-.41	-.21	.01	-.13	.17	-.15	.43
	% 111	.33	-.30	.48	-.12	-.38	-.32	-.02	-.07	.32	-.19	.36
Personal Control	f 101	.46	-.18	.54	-.13	-.31	-.33	.14	.12	.37	-.35	.56
	% 112	.39	-.19	.50	-.15	-.29	-.35	-.16	.11	.36	-.42	.51
Moralizing	f 102	.09	.06	.16	-.02	-.11	-.11	.20	.20	.16	-.25	.30
	% 113	.03	.05	.13	.03	-.09	-.11	-.16	.24	.16	-.27	.20
Hostility & Reprimands	f 103	.52	-.26	.59	.04	-.45	-.19	.03	.06	.43	-.20	.64
	% 114	.48	-.23	.55	-.02	-.40	-.19	.03	.04	.43	-.22	.60
Ignores Child	f 104	.03	-.16	.19	.09	-.12	-.09	-.04	.14	.20	-.10	.21
	% 115	-.01	-.12	.19	.05	-.08	-.09	-.09	.20	.19	-.13	.17

control variable tend to be slightly higher than in the first grade, particularly correlations with the punitive and controlling dimensions and also (negatively) with the supportive Q-sort. In contrast the personal control correlations tend to be somewhat lower though paralleling those in the first grade. The 'moralizing' category shows very low correlations with all Q-sorts for the third grade. The 'hostility and reprimands' scores correlate in a very similar pattern to that found in the first grade though at a somewhat lower level, most correlations being of the magnitude of .45 as opposed to .65 or higher in the first grade. Lastly the observation, 'ignoring child,' shows very low correlations with all Q-sorts.

In general then, the same patterns of correlations are found for both grades, the magnitude of correlation being generally somewhat lower in the third grade. In approaching the matrix from the opposite point of view and, for example, asking what observational data seem to be most important in relating to the Q-sorts the 'controlling' Q-sort correlates most highly with the 'personal control' and 'hostility and reprimands' categories as was true in grade one. The 'warmth' Q-sort correlates most highly with the 'praise and encouragement' observations, again as was true in grade one, and at a somewhat lower level with the 'non-verbal affiliation' and negatively with the 'personal control' categories. For the third grade the 'punitive' Q-sort correlates most highly with the 'personal control' and 'hostility and reprimands' categories but the correlations with 'academic control' and (negatively) with 'minimal reinforcements are of almost equal magnitude whereas in the first grade these variables correlated at a considerably lower level. The Q-sort on 'confidence' is virtually unrelated to the categories of observation in third grade, the one possible exception being a fairly low order correlation with 'problem structuring.' The 'supportive' Q-sort correlates positively with the 'praise and encouragement' and (negatively) with the 'hostility and reprimands' categories and negatively with both 'personal control' and 'academic control.' However, in the third grade all of these components are more equally represented whereas in the first grade the 'praise and encouragement,' 'hostility and reprimands,' and 'impersonal control' categories showed considerably higher correlations than the other variables. The 'stimulating' Q-sort is related most highly to the 'praise and encouragement' and 'personal control' (negatively) categories. This represents something of a change since for the first grade the 'hostility and reprimands' category was also negatively related to the Q-sort on 'stimulating' and this correlation drops considerably on the third grade. The Q-sort for 'achievement orientation' in the third grade is insignificantly related to any of the observations; whereas in the first grade it had shown significant correlations with the 'asking questions' category. Q-sort, 'physical contact,' again is primarily related to the 'non-verbal affiliation' observation as would be expected. 'Intellectual effectiveness' in the third grade correlates at a fairly low level with all of the categories, the highest being correlations with 'personal control' of the magnitude of -.36. In the first grade correlations of the magnitude of .50 were found with the 'asking questions' category. The last Q-sort, 'disparaging,' shows a very similar pattern to the first grade, the highest correlations (of the magnitude .50 to .65) found with 'personal control' and with 'hostility and reprimands.' Once again the overall finding is fairly good consistency with variables presumed to be similar with both techniques but the magnitude of correlations being generally lower in the third grade than in the first.

Interrelations Among Q-sorts.--As shown in Table 3.15, the pattern of intercorrelations among Q-sorts is much the same for the two grades. As might have been expected certain of the scales are fairly highly correlated; e.g., the 'warmth' scale is quite highly correlated with the 'supportive,' 'disparaging,' and 'punitive' scales; the latter two in a negative direction. This suggests that these four scales are measuring very similar variables. It is of considerable interest, however, to note that a second apparent factor is unrelated to these variables and this appears to be a factor loading on 'achievement orientation,' 'stimulating,' and 'intellectually effective.' It would appear that these three scales are measuring rather similar variables also. As would be expected, 'physical contact' correlates moderately with both the 'warmth' and 'supportive' scales. The correlations, however, are low enough (approximately .45) to indicate that the scales are far from synonymous. It is of some interest to note that the 'confidence' scale is quite unrelated to the 'warmth' dimensions and infact shows its highest correlations with the 'stimulating,' 'intellectually effective' and 'achievement oriented' scales; all of these being of the magnitude .30 to .40. Finally, and of considerable interest, is the variable 'controlling;' this being essentially the only scale showing differences between the first and third grade. In grade three this scale correlates very highly with the 'punitive' and 'disparaging' scales (.80 and .76 respectively) and only slightly lower with the 'supportive' and 'warmth' scales; the latter two being -.71 and -.60, demonstrating that in the eyes of the observers the teacher who is viewed as very controlling is also viewed as non-warm and quite punitive in the third grade. Within the first grade the scale seems

TABLE 3.15

INTERCORRELATIONS AMONG Q-SORTS (JOINT)--GRADES I AND III*

	Warm	Pun.	Conf.	Supp.	Stim.	Ach. Or.	Phys. Cont.	Intell. Effect.	Disp.
Controlling	-.39 (-.60)	.71 (.80)	-.01 (.20)	-.42 (-.71)	-.03 (-.26)	.22 (.33)	-.03 (-.23)	.02 (-.06)	.63 (.76)
Warm		-.74 (-.61)	-.11 (-.19)	.82 (.86)	.41 (.53)	-.24 (-.14)	.43 (.42)	.03 (.24)	-.64 (-.62)
Punitive			-.02 (.12)	-.76 (-.71)	-.26 (-.33)	.28 (.24)	-.22 (-.09)	.05 (-.17)	.80 (.84)
Confident				-.04 (-.13)	.33 (.29)	.45 (.44)	-.10 (-.17)	.27 (.47)	-.12 (.21)
Supportive					.53 (.55)	-.07 (-.12)	.34 (.42)	.23 (.29)	-.62 (-.71)
Stimulating						.57 (.40)	.07 (.11)	.68 (.70)	-.42 (-.44)
Achievement Oriented							-.00 (-.32)	.82 (.67)	.06 (.14)
Physical Contact								-.06 (-.16)	-.13 (-.18)
Intellectually Effective									-.10 (-.26)

*Note--Grade III in parentheses.

to be somewhat more independent of the 'warmth' and 'supportive' scales; the magnitude of correlations being somewhat lower. In third grade control is even more a function of the 'punitive' and 'disparaging' scales. This finding once again supports one of the findings of our previous research in that, in spite of the fact that we have attempted to define the dimensions of warmth, control and punitiveness independently and have tried to utilize these definitions in our observations, it turns out that for the most part the teacher who is viewed as the most controlling is also viewed as the most punitive and least warm.

CHAPTER IV

CORRELATIONAL ANALYSES

When faced with the approximately ten thousand intercorrelations resulting from a matrix of the size here used, it is obvious that some procedure must be established for study of these results. The following approach was undertaken within each of the two grades. The particular areas of interest were subdivided as follows:

1. Correlations between the test data obtained on the teachers and the observational measures of classroom behavior.

2. Correlations between the measures of pupil behavior and the observational data on teachers.

3. Correlations between the test data on teachers and the pupil data. Within each of these subdivisions the following strategy was employed. A number of hypotheses deriving primarily from previous research were tested. Thus, those measures pertaining to the particular hypothesis were scrutinized and reported and the evidence supporting or opposing the hypothesis is discussed. Clearly, however, there are many blocks of data wherein such hypotheses were not developed. For these parts of the data the following procedure was followed. Each variable was studied across the variables with which it had been correlated to determine the number of significant correlations which emerged. For both grades a standard error of 'r' of .16 was utilized since the sample size in both grades was between 36 and 40. Thus, a two-tailed test of significance at the 5 per cent level required a

correlation of .32. Whenever the number of correlations exceeding .32 was greater than chance and/or the correlations formed a coherent pattern, they are reported. It should be noted that the first prerequisite to interpretation of this kind is that the number of correlations in the matrix exceed those to be expected by chance. This is clearly the case. Although the total number of significant correlations, i.e., above .32, was not computed for the matrix, it is clear that for many of the variables of particular interest, the number of correlations with other meaningful variables which are significant greatly exceeds that to be expected by chance.

Relationships Between Teacher Test Measures and Teacher Classroom Behavior

It is of considerable interest to determine the extent to which meaningful categories of teacher classroom behavior are predictable from various test devices. Our earlier research (Travers and Wallen, 1961) had indicated quite strongly that one particular measure, the Teacher Preference Schedule, and in particular one score, that for Control Need, was quite an adequate predictor of the control--affiliation dimension of teacher behavior. Thus, in two separate samples correlations of the magnitude of .50 were obtained with measures of these dimensions. Consequently the first hypothesis investigated was the correlation between the TPS control score and the controlling and affiliation measures of classroom behavior. As shown in Table 4.01 this hypothesis receives only very weak support from the present data. Thus, of the 26 predictions made for the first grade, 22 out of 26 are in the correct direction as are 20 out of 26 in the third grade which is considerably above chance expectations. Further, the results are consistent

TABLE 4.01

PREDICTED CORRELATIONS BETWEEN TPS CONTROL AND MEASURES OF TEACHER BEHAVIOR

	TPS Control Grade I	 Grade III
Q-Sort: Controlling (Total)	.14	.25*
Q-Sort: Controlling (Joint)	.15	.24*
Q-Sort: Warm (Total)	-.08	-.33*
Q-Sort: Warm (Joint)	-.27*	-.16
Q-Sort: Punitive (Total)	.10	.18
Q-Sort: Punitive (Joint)	.15	.13
Q-Sort: Supportive (Total	-.18	-.22
Q-Sort: Supportive (Joint)	-.32*	-.28*
Q-Sort: Factor I (Total)	-.14	-.28*
Q-Sort: Factor I (Joint)	-.26*	-.18
Q-Sort: Physical Contact (Total)	-.15	-.35*
Q-Sort: Physical Contact (Joint)	-.23	-.27*
Observations: Praise and Encouragement (f)	-.21	.35*
Observations: Praise and Encouragement (%)	-.09	.36*
Observations: Non-Verbal Affiliation (f)	-.09	-.32*
Observations: Non-Verbal Affiliation (%)	-.04	-.27*
Observations: Academic Control (f)	-.15	.22
Observations: Academic Control (%)	.00	.16
Observations: Personal Control (f)	.09	-.10
Observations: Personal Control (%)	.20	-.07
Observations: Hostility (f)	.12	.08
Observations: Hostility (%)	.23	.09

TABLE 4.01 (Continued)

	Grade I	TPS Control Grade III
Ratings: Controlling	.09	.06
Ratings: Disparaging	-.19	-.07
Ratings: Supportive	.15	.03
Ratings: Affiliative	-.28*	-.19

* = Significant .05, 1 tail.

across the two grades with the one exception being the observations on 'praise and encouragement' where the correlation is negative in the first grade and positive in the third grade. The magnitude of the correlations, however, is much lower than hypothesized. Thus, in the first grade only 4 of the 24 are significant at the 5 per cent level using a one-tailed test; in the third grade 11 out of 26 are significant at the 5 per cent level using a one-tailed test. One would anticipate on the basis of chance that slightly more than one out of the 26 correlations would be significant by chance. In the first grade the significant correlations all in the expected direction are -.27 with Q-Sort (joint) for warmth, -.32 with Q-Sort supportive (joint), -.28 with rating affiliation, and -.26 with Q-sort Factor I (joint). For grade three the significant correlations are .25 with Q-sort controlling (total), .24 with Q-Sort controlling (joint), -.33 with Q-Sort warmth (total), .35 with observations 'praise and encouragement' (frequency score), .36 with observations 'praise and encouragement' (frequency score), .36 with observations 'praise and encouragement' (percentage score), -.32 with observations non-verbal affiliation (frequency score), -.27 with observations on verbal affiliation (percentage score), -.28 with Q-sort Factor I (total), and -.35 with Q-sort physical contact (total), and -.27 with Q-sort physical contact (joint). All of the significant correlations are in the hypothesized direction with the exception of the correlations with the observational measure of 'praise and encouragement' in the third grade. In total then the TPS control score does provide a set of correlations which exceed chance expectations in terms of predicted direction and a considerably greater than chance expectation as to the number of significant correlations, this being particularly true in the third grade. The magnitude of the correlations, however, is considerably smaller than had been anticipated on the basis of our previous studies, the highest correlations being of the magnitude of .35 and one of the major predictions in grade three showed a correlation opposite in direction to the prediction.

Next our attention moves to consideration of all of the teacher test measures employed with regard to their ability to predict teacher behavior. It would have been possible to set up a great many hypotheses here also, hypothesizing that tests intended to measure a particular variable would correlate with teacher behaviors having similar designation. Thus, one would expect TPS dominance to correlate with controlling behavior on the part of the teacher. However, rather than setting up a large number of specific hypotheses justifying a one-tailed test, it was decided instead to utilize two-tailed tests and to proceed by scrutinizing each of the test or predictor variables to determine the number of significant correlations with measures of teacher behavior.

The total number of behavioral measures considered was 54. One might argue that a smaller number than this should be used since 22 of the measures pertained to the Q-sorts wherein both total and joint sorts are used and the correlation between the total and joint sort is typically quite high. Further, the Q-sort and rating measures are quite highly related. Nevertheless, it was decided to treat these as separate measures since there was always the possibility, which infact occurred, that one of the scales might correlate whereas another did not. In any case, out of the total of 54, at the 5 percent level, one would expect three correlations by chance. Applying the standard error of proportion one would expect as many as six correlations to occur about 5 percent of the time by chance. Accordingly any predictor with seven significant correlations with teacher behavior measures was considered further. Table 4.02 indicates the number of significant correlations with teacher behavior measures for each of the test measures and also the biographical data available on the teachers. The distribution of numbers of significant correlations shown at the bottom of the table is interesting in that the distributions for the lower numbers of correlations (the top part of the table) seems to correspond reasonably well to the normal curve expectation. However, in both grades there are a number of radical departures from this to be found at the high frequency level. Thus, in grade one there are a total of six predictors showing 10 or more significant correlations and seven in grade three. It is of further interest to note that a number of the tests are to be found producing a large number of significant correlations in both grades. Among these are the EPPS Deference Scale, the EPPS Dominance Scale, and the EPPS Autonomy and TPS Child's Autonomy. Table 4.03 presents those

TABLE 4.02

NUMBER OF SIGNIFICANT CORRELATIONS OF TEACHER TEST MEASURES WITH TEACHER BEHAVIOR MEASURES

Test	Grade I	Grade III
TPS Achievement	1	15
TPS Affiliation	1	6
TPS Recognition	3	1
TPS Control	5	5
TPS Instrumental Rewards	2	5
TPS Status Striving	2	3
TPS Child's Affection	0	7
TPS Child's Autonomy	11	6
TPS Rebellious Motives	2	8
TPS Vicarious Motives	2	5
TPS Obsessive	1	11
TPS Dependency	2	12
TPS Exhibitionism	5	4
TPS Dominance	0	6
Objectives Achievement	2	0
Objectives Affiliation	3	2
Objectives Recognition	2	1
Objectives Control	2	1
EPPS Achievement	3	0
EPPS Order	3	6
EPPS Autonomy	12	6
EPPS Intraception	1	15

TABLE 4.02 (Continued)

Test	Grade I	Grade III
EPPS Dominance	13	9
EPPS Nurturance	1	7
EPPS Endurance	3	2
EPPS Aggression	7	5
EPPS Deference	10	10
EPPS Exhibition	5	0
EPPS Affiliation	3	0
EPPS Succorance	3	1
EPPS Abasement	1	0
EPPS Change	1	4
EPPS Heterosexuality	3	7
EPPS Consistency	4	7
Situations Achievement	0	7
Situations Affiliation	5	0
Situations Recognition	4	1
Situations Control	3	18
H-T-P Control	2	5
H-T-P Affiliation	2	3
H-T-P Achievement	0	0
H-T-P Recognition	5	0
H-T-P Ego Strength	0	1
PSI Control	12	0
PSI Affiliation #1	4	5

TABLE 4.02 (Continued)

Test	Grade I	Grade III
PSI Affiliation #2	3	3
Age	3	0
Years Taught	1	0
College	11	13
Degree	4	5
Hours Past A.B.	0	2
Years of Degree	5	3
College Major	7	3
Distribution:		
0	6	11
1	8	6
2	10	3
3	11	5
4	4	2
5	6	7
6	0	5
7	2	5
8	0	1
9	0	1
10	1	1
11	2	1
12	2	1
13	1	1
14	0	0
15	0	2
16	0	0
17	0	0
18	0	1

TABLE 4.03

CORRELATIONS BETWEEN THE BEST PREDICTORS OF TEACHER BEHAVIOR AND THE TEACHER BEHAVIOR MEASURES[1]

Behavior Measures	GRADE I						GRADE III										
	TPS Ch.	EPPS Aut.	EPPS Dom.	EPPS Agg.	EPPS Def.	PSI Con.	TPS Ach.	TPS Ch.	TPS Ch.	TFS Reb.	TPS Obs.	TPS Dep.	EPPS Intra.	EPPS Dom.	EPPS Nur.	EPPS Def.	Sit. Con.
Q-Sort Control		.26	-.21		.21	.28	.30				.36	.34	-.34	-.21	-.24	-.33	.32*
Q-Sort Control			-.24		.49*	.31	.32				.30	.23	-.44	-.25	-.22	-.37	.30
Q-Sort Warmth		-.41*	.38*		-.28	-.32*	.22		.21		-.25	-.21	.41	.34	.24	.23	-.40*
Q-Sort Warmth		-.34*	.39*		-.28	-.41*	-.28			-.23	-.22	-.27	.37	.30			-.28
Q-Sort Punitive		.33*	-.30		.22	.43*	.32			.27	.43	.39	-.25		-.35	-.25	.32*
Q-Sort Punitive		.28	-.29		.34*	.48*	.33				.37	.36	-.43		-.29	-.29	.22
Q-Sort Confident			.25	.42*	-.34*			-.25					-.28				
Q-Sort Confident				.32*	-.23								-.29				.26
Q-Sort Supportive	.22	-.22	.31		-.25	-.33*	-.27			-.20	-.26	-.25	-.30	.34			-.37*
Q-Sort Supportive	.25	-.32*	.33*	.20	-.22	-.34*	-.31			-.20	-.36	-.33	.43	.28	.22	.28	-.40*
Q-Sort Stimulating	.34*	-.33*						-.42		-.29	-.23					-.27	-.25
Q-Sort Stimulating	.35*		.25	.29	-.22			-.32			-.20			.25			
Q-Sort Achievement Oriented	.32*			.25				-.23				.25	-.37		-.23	-.55	
Q-Sort Achievement Oriented	.25			.30		.28		-.26					-.28		-.20	-.44	
Q-Sort Physical Contact	.23	-.30	.39*						.24				.32			.39	
Q-Sort Physical Contact	.22	-.32*	.37*										.24	.20		.38	-.38*
Q-Sort Intellectual Effectiveness	.34*			.24	-.25			-.35								-.33	
Q-Sort Intellectual Effectiveness	.27			.36*		.26		-.26									
Q-Sort Disparaging		.22			.30	.39*					.24	.23	-.46		-.28	-.31	.25
Q-Sort Disparaging		.20	-.38*		.54*	.34*	.30				.29	.25	-.46	-.26		-.31	.25
Q-Sort Factor I		.34	-.34*		.27	.38*	.32			.21	.38	.35	-.37	-.29	-.30	-.27	.40*
Q-Sort Factor I		.29	-.37*		.39*	.45*	.35				.34	.35	-.47	-.25	-.22	-.31	.32*
Observations Raised Hand (f)				.39*				-.25	-.48	-.33	-.23	-.44		-.45			
Observations Raised Hand (%)				.35*			-.24	-.22	-.45	-.32	-.31	-.50		-.37			-.20
Observations Praise and Encouragement (f)		-.31															
Observations Praise and Encouragement (%)		-.33*	.24					.20							.24		
Observations Non-verbal Affiliation (f)		-.27	.27						.26					.36	-.26		-.37*

Observations Raised Hand (f)				.39*				-.25	-.48	-.33	-.23	-.44		-.45			
Observations Raised Hand (%)				.35*			-.24	-.22	-.45	-.32	-.31	-.50		-.37			-.20
Observations Praise and Encouragement (f)		-.31															
Observations Praise and Encouragement (%)		-.33*	.24					.20							.24		
Observations Non-verbal Affiliation (f)		-.27	.27						.26					.36	-.26		-.37*
Observations Non-verbal Affiliation (%)	.21	-.21	.25			-.24			.20					.26	-.30		-.29
Observations Minimal Reinforcement (f)	.25			.33*	-.27				-.21				.22			-.23	
Observations Minimal Reinforcement (%)				.28	-.39*								.23	.28		-.32	
Observations Asking Questions (f)	.25			.26				-.24		-.29				-.26			-.22
Observations Asking Questions (%)	.29									-.34					.22		-.36*
Observations Problem Structuring (f)					-.35*				-.37								
Observations Problem Structuring (%)		.23				-.26	-.32		-.26							.24	
Observations Academic Control (f)		-.24					.35			.26	.35	.45					.33*
Observations Academic Control (%)		-.25					.40		.20	.39	.40	.61					.29
Observations Personal Control (f)					.25	.24	.21		.38	.24	.20	.25					.34*
Observations Personal Control (%)	-.35*			-.26	.20		.23		.38	.34	.23	.27					.38*
Observations Moralizing (f)					.26												
Observations Moraliz					.29		-.20					-.21			-.30		
Observations Hostility and Reprimands (f)		.39*	-.43*		.32*	.36*			.22	.30	.30	.23	-.26		-.28	-.25	.41*
Observations Hostility and Reprimands (%)	-.38*	.28	-.24	-.23	.23	.20			.20	.34	.26	.23	-.25		-.30	-.25	.43*
Observations Ignores Child (f)	-.45*	.23			.31				.26			-.21	-.24		-.36		
Observations Ignores Child (%)	-.42*	.24			.29				.29				-.23		-.33		
Rating Controlling					.37*		.20				.29		-.30	-.20	-.43		.32*
Rating Stimulating	.35*			.36*			-.37	-.42	-.26	-.41	-.39	-.28					-.27
Rating Disparaging			-.26		.28	[illegible]	.25			.22	.28	.26	-.28	-.33	-.21		.36*
Rating Supportive		-.38*	.35*		-.23	-.25				-.21			.23	.35	.21		-.43*
Rating Confident			.30	.28	-.22					-.26			-.21				
Rating Affiliative	.34*	-.37*	.26			-.26	-.26				-.22	-.25	.26	.47			-.30
Rating Relaxed	-.46*	.22		-.30				.42	.20								

---

[1] Correlations below .20 are omitted. Correlations of .32 and above sig. .05 level (two tail) are indicated by asterisk.

predictors showing large numbers of significant correlations with teacher behavior in the two grades along with the behaviors to which they relate. Also included are measures having only five or six significant correlations with teacher behavior but where those correlations themselves were quite high--several of them of magnitude of .40 or higher. Upon studying Table 4.03 it becomes apparent that certain predictive patterns of teacher behavior which are logically coherent emerge. Further, in several cases these are clearly consistent with the hypothesized variable being measured by the test. In other cases this latter correspondence is not so clear.

Grade I.--Overall the best predictor in grade one appears to be the PSI control score. Inasmuch as the findings of our prior and present research demonstrate a general tendency for negative correlations between controlling dimensions and affiliative type dimensions, it is to be expected that a predictor relating to one of these dimensions may relate to the others as was the case in our prior research with the TPS control score. In the present instance, the PSI control score correlates significantly in the expected direction with 17 out of 53, or 31 per cent, of the measures of teacher behavior if one were to utilize a one-tailed test and the 5 per cent level. Further, four of these correlations are .40 or larger, suggesting practical as well as theoretical use. Thus, there appears to be considerable support in Table 4.03 for the following generalizations about the first grade teacher scoring high on the PSI control score. She is very likely to be viewed as less warm and less supportive; more disparaging and more punitive by observers and to have a high frequency in the category of observed 'hostility and reprimands' exhibited in her classes. There is somewhat less support in terms of the magnitude of correlations but nevertheless considerable support across several indices to indicate that she is, in addition, more controlling, and exhibits less non-verbal affiliation toward her pupils. In addition, there is a weaker suggestion that she is more achievement oriented, and exhibits less problem structuring behavior.

It will be noted that quite similar patterns of correlations emerge with three of the four EPPS scales shown in Table 4.03; the Autonomy, Dominance, and Deference scales, though in the case of the Dominance scale, the correlations are in the opposite direction of those with the Autonomy and Deference and PSI Control scores. Thus, these three scales also seem to predict fairly well the teacher variables of controlling, warmth, punitiveness, supportive behavior as viewed by the observers and also the objective categorizing of 'hostility and reprimands.' The principle differences in predictive power of these scales, however, are that the Autonomy scale seems to have some possibilities for predicting stimulating teacher behavior as well as the specific category of 'praise and encouragement' as viewed objectively. Further, the Autonomy and Dominance scales predict extent of physical contact better than do the other two scales. For purely predictive purposes, therefore, there seems to be relatively little to choose among these four scales although the correlations tend to be somewhat higher for the EPPS deference and PSI control scales.

From a theoretical point of view two of these EPPS scales are rather perplexing. Thus, it is probably theoretically to be expected that teachers scoring high on Autonomy would tend to show the pattern which emerges, i.e., being more controlling, more punitive, showing less warmth, less supportive behavior, less physical contact, less praise

TABLE 4.04

INTERCORRELATIONS AMONG THE BEST PREDICTORS OF TEACHER BEHAVIOR--GRADE I

	EPPS Autonomy	EPPS Dominance	EPPS Aggression	EPPS Deference	PSI Control
TPS Child's Autonomy	.00	-.09	.13	-.11	-.09
EPPS Autonomy		-.34	.26	-.16	.27
EPPS Dominance			-.01	-.28	-.45
EPPS Aggression				-.21	.11
EPPS Deference					.34

and encouragement, less non-verbal affiliation, more hostility, etc., if one assumes that such a need predisposes the teacher to wish to remain aloof from her children and preserve her individual autonomy. However, it is difficult to understand the direction of the EPPS Dominance and Deference correlations. Thus, the correlations for the Dominance scale are generally the reverse of what might be expected, in that it might be hypothesized that a teacher high in dominance need would be more controlling, less warm, more punitive, less supportive, etc., whereas in fact the Dominance scale correlates in the opposite direction with all of these variables. Further, the pattern of direction of correlations with the Deference scale is also surprising in that one might have expected the more deference teacher to be somewhat less controlling, more warm, less punitive, etc., whereas, in fact, these correlations are in the reverse order.

Of the two remaining predictors the TPS Child's Autonomy score shows promise for predicting certain aspects of teacher behavior not well predicted by the preceding four scales, particularly stimulating behavior, degree of achievement orientation, intellectual effectiveness as observed, degree of personal control exerted, and ignoring the child, as well as the intensity vs. relaxed dimension. Most of these correlations make theoretical sense in that one would expect to find that the teacher being more concerned with the Child's Autonomy and respecting the child would be more supportive, engage in more physical contact, be less controlling of personal behavior, show less hostility, and ignore the child to a lesser extent. Theroetically, however, there seem to be no good grounds for expecting some of the other higher correlations, e.g., the teacher high in 'child autonomy' is viewed as being more intense, more achievement oriented, and more stimulating.

The remaining predictor shown for the first grade, EPPS Aggression adds relatively little to the prediction already possible. It does predict the extent of 'confident' behavior as viewed better than other measures and also is the only predictor to predict with any success the extent to which the teacher acknowledges the children's raised hands.

It is of interest to examine the intercorrelations among the six predictors within the first grade particularly since four of them seem to be tapping very similar aspects of teacher behavior. These correlations are presented in Table 4.04. It will be noted that the TPS Child's Autonomy score which predicted a somewhat different domain of teacher behavior from the others shows very low correlations with the other five tests. The EPPS Aggression score which added predictive ability in one or two instances tends to show rather low correlations with the other scores.

As would have been expected, the remaining four tests, EPPS Autonomy, EPPS Dominance, EPPS Deference, and PSI Control tend to show somewhat higher correlations among themselves and in directions consistent with their correlations with teacher behavior. Thus, PSI control score is positively correlated (.27 and .34) with the EPPS Autonomy and EPPS Deference. However, the correlation between EPPS Deference and EPPS Autonomy is surprisingly negative though low (-.16). Further, EPPS Dominance which tended to predict the same pattern of teacher behaviors as the other scales but with reversed signs does correlate negatively from -.28 to -.45 with the other three scales. Clearly then there is a consistency in the way in which these tests relate among themselves as well as how they predict teacher behavior. However, the correlations are for the most part not high and suggest that the tests themselves are getting at somewhat different teacher characteristics although there are considerable similarities as to the teacher behavior which is predicted.

Grade III.--The best predictor of teacher behavior in grade three turns out to be the Situations Control Measure which has significant correlations with 18 indicators of teacher behavior and the nature of these relationships is quite consistent. Thus, we find significant positive correlations with the Q-sort measure of controlling behavior, Q-sort measure of punitive behavior, Q-sort measure of disparaging behavior, and the Factor I Q-sort indicating controlling, non-permissive behavior and with the rating control, rating disparaging measures. Next we find significant negative correlations with the Q-sort measures of warmth, supportive behavior, and with the ratings of supportive behavior, affiliative behavior. With the observational categories we find significant positive correlations with both 'academic control' and 'personal control' and with frequency of 'hostility and reprimands' and negative correlations with 'non-verbal affiliation.' In addition to thise consistent pattern we also find negative correlations with both the Q-sorts and ratings on 'stimulating' and for the observational category 'asking questions.' It is of considerable interest to note the very striking similarity of the predictions made by this test for the third grade with those made by the PSI control score in the first grade. One would hope, of course, that the same instrument might function in a similar fashion in the two grades. What is suggested, however, is that a more subtle projective device predicts better in the first grade whereas a more situationally oriented test predicts better in the third grade. It is likely that the situations portrayed in the latter test fit the third grade better than the first.

As shown in Table 4.03, the second best predictor of teacher behavior in grade three is the EPPS Intraception scale, which shows significant correlations with a number of the observer judgments and near significant correlations with a number of the observational category scores. In general, the teacher scoring high on the Intraception scale is viewed as less controlling, more warm, less punitive, somewhat less confident, more supportive, less achievement oriented and less disparaging. Three of the other Edwards scales: Dominance, Nurturance, and Deference also seem to predict in a similar fashion to that of the Intraception scale, the only differences of importance being that the Nurturance scale shows somewhat higher correlations with the objective category scores and somewhat lower correlations with the observer impressions and that the Deference scale, though generally predicting somewhat more poorly than the

Intraception scale appears to be a better predictor of achievement orientation showing sizeable correlations -.55 and -.44 with the observer impressions as to the degree of achievement orientation of the teacher. For the most part these scales make theoretical sense in that one would have expected the teacher scoring higher on Deference, Nurturance, and Intraception, to be generally the more permissive, warm type teacher. The exception is the Dominance scale which correlates in the same direction as the other scalesthough one might have expected correlations in the opposite direction. It should be noted, however, that the same finding was observed in grade one. The Dominance scale on the Edwards is functioning in grade three in a very similar fashion to its functioning in grade one, though it predicts at a somewhat poorer level in grade three. By contrast the EPPS Deference scale is functioning in the opposite direction from its functioning in grade one; that is, whereas in grade three a high score on the Deference scale indicates less disparaging behavior and less objectively recorded hostility; in the first grade the reverse was true and the teachers scoring higher on the Deference scale were viewed as more controlling, less warm and more punitive.

Of the other scales for the third grade the TPS Obsessive and TPS Dependency scales seem to be functioning in a similar manner, both showing positive correlations of considerable magnitude with observer impressions of punitiveness on the part of the teacher which is supported by positive correlations with extent of control over the student, the correlations being higher with 'academic control' than 'personal control.' The highest cluster of correlations to be found anywhere in Table 4.03 appears with both of these measures in predicting extent of academic control where the correlations range from .35 to .61. In this instance these findings are somewhat unexpected from a theoretical standpoint in that it is not clear why the more dependent and obsessive teacher would be the more punitive and more academically controlling teacher. Of the other measures the TPS Rebellious Motive score seems to be functioning in much the same way as the Obsessive and Dependency scores but generally predicting at a somewhat lower level. The TPS Achievement score appears to predict primarily the controlling-warmth dimension though less efficiently for the most part than tests previously mentioned. It tends to show positive correlations with various indices of controlling behavior and negative correlations with indices of supportive affiliative behavior.

The TPS Child's Autonomy scale shows substantial correlations with only two dimensions of classroom behavior. It is a good predictor of the teacher's tendency to react to children's raised hands, correlating negatively, and also correlates .38 with the objectively scores incidence of personal control. These findings do not make theoretical sense since it is not clear why a teacher high in need to allow children to have autonomy should be less reactive to the children's hands and also exert more personal control. The remaining scale, TPS Child's Affection, is of interest because it seems to predict in a somewhat different domain. It shows very low correlations with the dimensions of controlling and affiliative behavior but is the only test to show significant correlations with the indices of stimulating behavior on the part of the teacher and with observer impressions of intellectual effectiveness on the part of the teacher. The magnitude of these correlations is quite high, with correlations of -.42 with both the Q-sort and rating judgments of the 'stimulating' characteristics and .42 with a rating on relaxed behavior

Intercorrelations among these tests are shown in Table 4.05. In accordance with the previous discussion one would expect to find substantial correlations among TPS Achievement, TPS Dependency, TPS Rebellious, TPS Obsessive, and Situations Control and this is borne out. One would also expect correlations among EPPS Intraception, Dominance, Deference and Nurturance but this is not the case, possibly due to the difference in construction of the two tests, the EPPS being ipsative. As expected, TPS Child's Affection and Child's Autonomy appear to measure variables which are largely independent of the other measures.

Relationships Between Teacher Classroom Behavior and Pupil Measures

With regard to these relationships a number of specific hypotheses based on prior research were formulated. Within each domain of pupil data, these are discussed first. Subsequently, relationships not hypothesized are discussed. Within the tables in this section, those variables pertaining to hypotheses are labeled, e.g., H_1.

Pupil Achievement

Hypothesis I.--Positive correlations will be found between achievement gain, especially in reading vocabulary, and the extent to which the teacher was viewed as stimulating by observers. Predicted for both grades.

As shown in Table 4.06 this hypothesis is clearly supported though more strongly in grade one. The correlations are of the magnitude of .30 to .50 with end-of-year scores. When partial correlations were obtained resulting in correlations with achievement 'gain,' the first grade values remain essentially the same. In the third grade the correlations with Q-sort Stimulating (Total) reamin significant wheras

TABLE 4.05

INTERCORRELATIONS AMONG THE BEST PREDICTORS OF TEACHER BEHAVIOR--GRADE III

	TPS Achievement	TPS Rebellious	TPS Obsessive	TPS Dependency	TPS Intraception	TPS Dominance	TPS Nurturance	TPS Deference	TPS Child's Affection	TPS Child's Autonomy
Situations Control	.50	.53	.52	.46	.09	-.23	-.16	-.20	.22	-.09
TPS Achievement		.57	.75	.75	-.01	-.34	.00	-.02	.32	.19
TPS Rebellious			.54	.61	.16	-.03	-.13	-.08	.22	.18
TPS Obsessive				.67	-.02	-.18	-.12	-.16	.31	.13
TPS Dependency					.03	-.02	.01	-.12	.39	.34
EPPS Intraception						.24	.18	.02	.21	.14
EPPS Dominance							-.09	-.39	.01	.29
EPPS Nurturance								.17	.32	.02
EPPS Deference									-.01	-.20
TPS Child's Affection										.39

TABLE 4.06

CORRELATIONS BETWEEN CLASS ACHIEVEMENT AND MEASURES OF TEACHER BEHAVIOR

	Reading Vocabulary-Fall		Reading Vocabulary-Spring		Reading Comprehension-Spring		Arithmetic-Spring	
	Grade I	Grade III	Grade I	Grade III	Grade I	Grade III	Grade I	Grade III
Q-Sort Stimulating--Total H_1	-.02	.28	.28 (.293)*	.39 (.283)	.32 (.319)	.38 (.271)	.50 (.507)	.45 (.367)
Q-Sort Stimulating--Joint H_1	-.07	.37	.33 (.358)	.27 (.022)	.30 (.298)	.35 (.097)	.53 (.545)	.29 (.087)
Rating Stimulating H_1	-.23	.23	.32 (.397)	.31 (.215)	.29 (.287)	.24 (.096)	.32 (.363)	.13 (-.013)
Q-Sort Int. Effectiveness--Total	-.01	.48	.33 (.342)	.57 (.376)	.37 (.370)	.58 (.372)	.38 (.384)	.50 (.298)
Q-Sort Int. Effectiveness--Joint	-.05	.42	.38 (.404)	.45 (.244)	.48 (.479)	.56 (.411)	.34 (.349)	.54 (.395)
Observation Minimal Reinforcement (f)	.09	.18	.29 (.278)	.19 (.092)	.38 (.386)	.18 (.061)	.43 (.423)	.18 (.090)
Observation Minimal Reinforcement (%)	.26	.13	.17 (.110)	.04 (-.069)	.12 (.138)	.08 (-.040)	.39 (.372)	.19 (.141)
Observation Asking Questions (f)	-.11	.05	.40 (.439)	.13 (.132)	.53 (.528)	.11 (.117)	.23 (.248)	.25 (.277)
Observation Asking Questions (%)	.10	-.09	.41 (.401)	-.01 (.072)	.31 (.317)	.00 (.120)	.21 (.200)	.23 (.361)
Rating Confidence	.05	.09	.33 (.328)	.10 (.053)	.24 (.243)	.13 (.097)	.51 (.508)	.17 (.197)
Rating Relaxed	.12	-.21	-.36 (-.403)	-.45 (-.431)	-.47 (-.468)	-.39 (-.378)	-.35 (-.372)	-.26 (-.170)
Q-Sort Achievement Oriented--Total	.18	.30	.43 (.405)	.50 (.424)	.50 (.518)	.50 (.454)	.54 (.530)	.48 (.393)
Q-Sort Achievement Oriented--Joint	.08	.50	.48 (.476)	.55 (.327)	.56 (.567)	.61 (.404)	.43 (.425)	.52 (.313)
Q-Sort Physical Contact--Total	-.05	-.34	.25 (.270)	-.47 (-.346)	.24 (.239)	-.37 (-.174)	.34 (.350)	-.32 (-.151)
Q-Sort Physical Contact--Joint	-.03	-.24	.13 (.141)	-.51 (-.490)	.15 (-.149	-.38 (-.323)	.18 (.186)	-.31 (-.213)
Observations Personal Control (f)	.05	-.22	-.22 (-.239)	-.26 (-.153)	-.09 (-.088)	-.26 (-.144)	-.13 (-.137)	-.20 (-.085)
Observations Personal Control (%)	-.10	-.28	-.39 (-.379)	-.35 (-.226)	-.13 (-.136)	-.35 (-.219)	-.23 (-.220)	-.28 (-.144)

*Values in parenthesis are partial correlations, holding Reading Vocabulary Fall constant or 'in effect' correlations with achievement gain.

the others do not. Whether this decrease is due to the impact of the teacher's behavior being felt before the fall testing is a question which these data cannot answer.

In addition to the hypothesized relationships just discussed, Table 4.06 also indicates a number of other measures of teacher behavior which correlate significantly with achievement gain. For the most part these additional measures of teacher behavior fit the picture of the stimulating, achievement oriented, intellectually effective teacher being more successful in bringing about achievement gain in her class. Thus, the most substantial relationships are found with the Q-sorts on 'intellectually effectiveness' and 'achievement orientation.' These measures show correlations with achievement gain of the magnitude of .30 to .50 for both grades. It may be noted that in the first grade the correlation with achievement spring scores tends to be fairly high; the correlation with fall scores is nearly zero, as would be expected, resulting in a substantial correlation with achievement gain whereas in the third grade, although the correlations with spring score tend to be in many instances higher than those in the first grade, there are also in several instances rather sizeable correlations with achievement scores in the fall of the year. We tentatively attribute this to the impact which the teacher has had upon the pupils during the first six weeks or so of school before our testing took place. In any event the more important finding is that the correlations with gain during the year are significant in the third grade as well as the first.

Of the observations measures the major one relating to achievement gain is the 'asking questions' category which shows rather substantial correlations with achievement gain primarily in the first grade; in the third grade the only significant correlation is with arithmetic gain. Another finding applying only to the first grade is that the rating on 'confidence' is correlated significantly with achievement gain in reading vocabulary and arithmetic and at a somewhat lower level for comprehension. Again, in the first grade only, the observations category, 'minimal reinforcement,' shows several significant correlations with gain, primarily in arithmetic. The correlations with gain are positive also in the reading vocabulary and comprehension areas but only one of these is significant. The category, 'personal control,' shows negative correlations with all indices of achievement gain but only one of these reaches a significant level; this being a correlation of -.38 with gain in reading vocabulary in the first grade. Finally there is a rather interesting difference in the apparent impact of physical contact on achievement for the two grades, in that the Q-sort measure of physical contact shows positive correlations with all measures of achievement gain in the first grade though only one of these is significant at .35 with arithmetic gain. In the third grade, by contrast, all correlations are negative and three of these are significant ranging from -.32 to -.49 suggesting that physical contact in the first grade may foster achievement gain while hindering it in the third grade.

Pupil Anxiety

Hypothesis II.--Supportive behavior on the part of the teacher is negatively correlated with anxiety. Predicted for both grades.

As can be seen in Table 4.07 this hypothesis receives some support within the third grade only, in that the observational measure of praise and encouragement is significantly negatively correlated with test anxiety at the end of the year. The Q-sort measure of supportive behavior provides

TABLE 4.07

RELATIONSHIPS BETWEEN PUPIL QUESTIONNAIRE MEASURES OF ANXIETY AND TEACHER CHARACTERISTICS

	GRADE I			GRADE III		
	General Fall	General Spring	Test Anxiety Spring	General Fall	General Spring	Test Anxiety Spring
H_2 Q-Sort Supportive (Total)	.08	.18	-.06	.11	.13	-.20
H_2 Q-Sort Supportive (Joint)	.19	.13	-.21	-.03	.21	-.16
H_2 Rating Supportive	.18	.25	- 02	-.02	.05	.13
Q-Sort Confident (Total)	-.36	.00	-.20	.24	.02	.21
Q-Sort Confident (Joint)	-.44	-.24	-.34	.23	-.02	.25
Q-Sort Achievement Oriented (Total)	-.32	-.20	-.31	-.29	-.15	-.13
Q-Sort Achievement Oriented (Joint)	-.26	-.31	-.27	-.27	.00	-.06
Observations Answers Raised Hand (f)	-.05	.00	-.31	-.01	.17	.13
Observations Answers Raised Hand (%)	-.02	.17	-.33	-.02	.12	.05
H_3 Q-Sort Physical Contact (Total)	.00	.35	-.02	.19	.21	.17
H_3 Q-Sort Physical Contact (Joint)	.00	.26	-.16	.15	.04	.06
H_3 Observations Non-Verbal Affiliation(f)	.01	.37	-.06	.13	.07	-.03
H_3 Observations Non-Verbal Affiliation(%)	.12	.25	-.05	.09	.19	-.02
H_2 Observations Praise and Encouragement (f)	-.03	.17	.00	.11	-.08	-.31
H_2 Observations Praise and Encouragement (%)	.00	.36	.11	-.27	-.12	-.35

some additional support, both correlations being negative though not significant. The rating as to the extent of supportive behavior is exceedingly low and in the wrong direction .13. There is no support for this hypothesis within the first grade, only one of the measures being significant and this in the wrong direction (praise and encouragement).

Hypothesis III.--Overt-affiliative behavior on the part of the teacher is positively associated with anxiety. Predicted for the first grade only.

This hypothesis received considerable support from the data in that both the Q-sort measure of physical contact and the observational measure of non-verbal affiliative behavior correlate significantly with the general anxiety score in the spring, though correlating at a zero level in the fall, strongly supporting the hypothesis that the overtly affectionate teacher generates anxiety in first-graders. This hypothesis, though not specifically offered for the third grade, finds no support within that grade.

In addition to these specific hypotheses the correlations with anxiety were empirically examined according to the criteria previously stated Within the first grade considerably more than a chance number of correlations are significant with both general anxiety and test anxiety in the spring. In addition to the general anxiety correlates previously discussed, the significant correlations suggest that an achievement orientation on the part of the teacher is negatively correlated with general anxiety and also with test anxiety in the spring of the year. These correlations also exist, however, in the fall of the year and hence interpretation is difficult. Also, there is a suggestion that confidence on the part of the teacher is negatively correlated with both general anxiety and test anxiety in the spring but even more so in the fall, the correlation being -.44 in the fall; this being the highest correlation of any of the measures of teacher behavior with any of the anxiety measures, strongly suggesting that the confident teacher, within the early days of the first grade in particular, has a major impact resulting in less anxiety on the part of the students. As the years on this relationship still appears to hold though at a somewhat lesser level. The one correlation which is not expected nor consistent is the correlation of .36 between the general anxiety score in the spring and the observational score of praise and encouragement. It is not clear why this situation exists unless it once again indicates that the overtly-affectionate or perhaps even supportive behavior on the part of the first grade teacher generates anxiety in the first-grader. In the third grade there are no more than a chance number of significant correlations.

Liking for School

Hypothesis IV.--Liking for school will be negatively correlated with the observer measures of achievement orientation on the part of the teacher. Predicted for both grades.

Looking both at the Q-sort and rating measures of this variable it was discovered that there are no significant relationships in either grade. Hence, there is no support for this hypothesis. Note, however, that there is support for the hypothesis as applied to teacher personality as assessed through the test measures (discussed in the next section).

Hypothesis V.--Liking for school will be correlated with measures of warmth and permissiveness. Predicted for third grade only.

This hypothesis is strikingly confirmed by a great many measures as shown in Table 4.08. The overall pattern is clearly one of preference

TABLE 4.08

'LIKING FOR SCHOOL' RELATED TO TEACHER BEHAVIOR--GRADE III

Teacher Behavior	r
Q-Sort Controlling - Total H_5	-.24
Q-Sort Controlling - Joint H_5	-.25
Q-Sort Warmth - Total H_5	.48
Q-Sort-Warmth - Joint H_5	.41
Q-Sort Punitive - Total H_5	-.27
Q-Sort Punitive - Joint H_5	-.20
Q-Sort Supportive - Total H_5	.44
Q-Sort Supportive - Joint H_5	.43
Q-Sort Physical Contact - Total H_5	.35
Q-Sort Physical Contact - Joint H_5	.24
Q-Sort Disparaging - Total H_5	-.29
Q-Sort Disparaging - Joint H_5	-.29
Q-Sort Warm-Permissive - Total H_5	.40
Q-Sort Warm-Permissive - Total H_5	.36
Q-Sort Stimulating - Total H_5	.11
Q-Sort Stimulating - Joint H_5	.30
Observations - Answers Hand f	-.31
Observations - Answers Hand %	-.28
Observations - Non-verbal Affiliation f H_5	.39

TABLE 4.08 (Continued)

Teacher Behavior	r
Observations - Non-verbal Affiliation % H_5	.37
Rating - Disparaging H_5	-.34
Rating - Supportive H_5	.34
Rating - Relaxed	.50

for the warm friendly, supportive teacher and this dimension appears to be more important than the control dimension although several measures of control are consistent with the prediction though the magnitude is somewhat lower. Thus, it seems quite clear that by the third grade student's opinions of teachers whom they like are related to observer assessments of teacher behavior; most highly to dimensions that have to do with being relaxed, affiliative and friendly and to a lesser extent being unconcerned about control. Perhaps the surprising significant correlations are the negative correlations with the extent to which the teacher answers hands when raised. However, this may be consistent with the generally relaxed posture of such a teacher. It is of some interest in this connection to note the remaining measures of teacher behavior which were not related to teacher liking at a significant level. Thus, variables pertaining to confidence, achievement orientation, and intellectual effectiveness are in general unrelated to liking. This is reflected in the Q-sorts, ratings, and in the observational categories which pertain to reinforcement, asking questions, problem structuring, etc. In general these correlations are quite low. Further, it is interesting to note that the observations categories of 'hostility and reprimands,' 'academic control' and 'personal control' do not correlate significantly though they are in all cases negative as would be expected.

Hypothesis VI.--'Liking for school' will be positively correlated with measures of controlling behavior on the part of the teacher. Predicted for grade one only.

Hypothesis VII.--'Liking for school' will be negatively correlated with measures of affiliative behavior on the part of the teacher. Predicted for grade one only.

These hypotheses reflect one of the more interesting findings of our prior research to the effect that within the first grade the teachers viewed as less permissive, more controlling, and less affiliative seemed to be better liked by their pupils. Our present data appear to clarify the nature of this prior finding. First, the direct measures of affiliative and controlling behavior on the part of the teacher do not show significant correlations with liking for school. However, the Q-sort measure for physical contact on the part of the teacher does show negative correlations with liking for school, one of which is at a significant level -.29, utilizing a one-tailed test. Further, the observational measures of physical contact also show negative correlations (-.14 magnitude). Thus, it is our interpretation that it is the overt affiliative behavior, personal contact, etc. which results in the teacher within the first grade being less well liked. There is, however, no support for our hypothesis that the teacher who is more controlling and has a better structured classroom is better liked by the students.

Divergent Thinking

Hypothesis VIII.--The more controlling teacher will have a depressing effect on creativity or divergent thinking as measured by the Torrance Circles Test or conversely; the more permissive teacher fosters such thinking. Predicted for both grades.

This hypothesis receives very little support. There are no significant correlations with spring scores or gain in the first grade. In grade three, there is a significant negative correlation with observations--'personal control' (f) in the spring but not with gain. As shown in Tables 4.09 and 4.10, it appears that in grade three the extent of controlling behavior on the part of the teacher is a much less potent variable than the domain of intellectual stimulation and achievement orientation, in that gain on the Circles Test is significantly related to Q-sort measures of teacher stimulating, achievement orientation and intellectual effectiveness and also to the Observations--'problem structuring'--and negatively--'asking questions'--which we would interpret as suggesting that the nature of most questions asked is pretty rhetorical.

It will be recalled that two measured were used with the Circles Test; one the manual scoring and the other a rating. It will be noted that for grade three the two measures agree quite well in terms of their relations with other variables. This is to be expected since the correlation between the two scores is .46 in the fall and .58 in the spring. Thus, although the two scores are not measuring precisely in the same fashion, there is considerable agreement between them. When looking at the results of the first grade, however, several very different results become evident. First of all the two scores show virtually no agreement;

TABLE 4.09

CORRELATIONS BETWEEN 'CIRCLES' TEST OF DIVERGENT THINKING AND MEASURES OF TEACHER BEHAVIOR--GRADE I

	Score		Score	
	Fall	Spring	Fall	Spring
Observations Moralizing (f)	.03	-.04	.46	.13
Observations Moralizing (%)	-.33	-.17	.53	.31
Observations Hostility and Reprimands (f)	-.06	.06	.35	.02
Observations Hostility and Reprimands (%)	-.06	.03	.31	-.08
Observations Praise and Encouragement (f)	.27	.12	-.24	-.12
Observations Praise and Encouragement (%)	.44	.24	-.50	-.25
Observations Asking Questions (f)	.02	.05	-.05	-.18
Observations Asking Questions (%)	-.02	.11	-.37	-.30
Rating Controlling H_6	-.14	-.01	.43	.21
Rating Disparaging	-.08	.07	.41	.18
Rating Supportive	.08	.07	-.31	.07
Observations Minimal Reinforcement (f)	.32	.23	.12	.09
Observations Minimal Reinforcement (%)	.31	.25	-.08	.10
Observations Ignores Child (f)	-.44	-.13	.26	.22
Observations Ignores Child (%)	-.46	-.17	.24	.23
Observations Problem Structuring (f)	.17	-.30	-.21	-.21
Observations Problem Structuring (%)	-.43	-.43	.19	.33
Observations Academic Control (f) H_6	.26	.13	.12	.05
Observations Academic Control (%) H_6	.31	.11	-.14	-.09

TABLE 4.10

CORRELATIONS BETWEEN 'CIRCLES' TEST OF DIVERGENT THINKING AND MEASURES OF TEACHER BEHAVIOR--GRADE III

	Score		Rating	
	Fall	Spring	Fall	Spring
Q-Sort Stimulating (Total)	-.10	.35 (.410)	-.16	.17
Q-Sort Stimulating (Joint)	.13	.36 (.339)	.18	.40
Q-Sort Achievement Oriented (Total)	-.21	.31 (.414)	-.01	.21
Q-Sort Achievement Oriented (Joint)	-.15	.33 (.410)	.11	.24
Q-Sort Intellectual Effectiveness (Total)	-.04	.39 (.430)	.14	.31
Q-Sort Intellectual Effectiveness (Joint)	.13	.35 (.328)	.24	.33
Q-Sort Disparaging (Total)	-.29	-.16 (-.068)	-.30	-.16
Q-Sort Disparaging (Joint)	-.12	-.23 (-.202)	-.19	-.30
Observations Ignores Child (f)	-.31	.05 (.173)	-.07	.32
Observations Ignores Child (%)	-.24	.06 (.155)	-.06	.29
Observations Asking Questions (f)	.14	.00 (-.052)	-.13	-.18
Observations Asking Questions (%)	.22	-.02 (-.103)	-.24	-.30
Observations Problem Structuring (f)	.12	.23 (.202)	.16	.05
Observations Problem Structuring (%)	-.37	.21 (.385)	.06	.31
Rating Stimulating	-.15	.27 (.345)	-.08	.25
Rating Relaxed	.07	-.25 (-.292)	.22	-.01
Observations Personal Control (f) H_6	-.05	-.16 (-.152)	-.25	-.26
Observations Personal Control (%) H_6	.00	-.11 (-.117)	-.17	-.19

the fall correlations being -.17 and the spring .07. Thus, there is very little in common between these two measures in this grade. It will be recalled that our reason for resorting to the rating was our feeling that the manual score gave much too heavy weight to sheer repetition, redundancy or fluency. Thus, we are relying more heavily on the rating measure in our interpretation. Whether one looks, however, at the rating or the score measure within the first grade, several peculiarities become evident. The first is that the pattern found in the third grade clearly does not exist. Second, the number of correlations with teacher behavior is for each score, two in the spring; the number to be expected by chance. There are, however, much more than the number of chance correlations with the fall testing which is most peculiar. Thus, one cannot say much about teacher behavior which is likely to foster or hinder better performance on the Circles Test in the first grade. One can, however, explore the possible meaning of the fall correlations. In doing so we shall rely on the rating measure viewing it as the better index of divergent thinking. The picture then is such that the classes doing better on this measure in the fall are those whose teachers were viewed as high in moralizing, high in hostility and reprimands, low in praise and encouragement, low in asking questions, high in controlling, high in disparaging and low in supportive. This picture then is of a hostile, rigid, unfriendly teacher. Why this characteristic on the part of the teacher should relate to high scores on the Circles Test at the beginning of the year must remain something of a mystery. Perhaps this rather punitive introduction to school forces the child back on some kind of inner-resources which cause him to behave in a somewhat erratic free-associative fashion and perhaps this is what is revealed on the Circles Test.

Peer Ratings

The next pupil measure examined is the sociometric measure which was available only for the spring. It will be recalled that the measure used is the class mean rating given which is the same as the mean rating received for the class although these values differ for individuals. The assumption is made that the higher the mean score the more the pupils within the class view one another as attractive and as capable of satisfying their needs. In the first grade, more than a chance number of relationships with teacher behavior are found as shown in Table 4.11. The

TABLE 4.11

CORRELATIONS BETWEEN MEAN PEER EVALUATION BY CLASS AND MEASURES OF TEACHER BEHAVIOR--GRADE I

Q-Sort Intellectual Effectiveness--Total	.35
Q-Sort Intellectual Effectiveness--Joint	.38
Observations--Praise and Encouragement f	.05
Observations--Praise and Encouragement %	-.33
Observations--Asking Questions f	.36
Observations--Asking Questions %	.06
Observations--Academic Control f	-.02
Observations--Academic Control %	-.39
Observations--Moralizing f	.32
Observations--Moralizing %	.52
Observations--Ignores Child f	.32
Observations--Ignores Child %	.29

general picture which emerges from this table is that higher mean sociometric choice tends to result when the teacher is intellectually effective and prone to moralizing while at the same time ignoring the child to some extent. Further, there is the suggestion that 'praise and encouragement' and 'academic control' are both negatively related to this variable, whereas 'asking questions' is positively related. Our interpretation of these findings is to the effect that the moralizing which the teacher exhibits in the first grade which is frequently directed toward socially acceptable behavior is effective since the highest correlation (.52) is found with the per cent of moralizing behavior found on the part of the teacher. In addition, there is a suggestion that praise and encouragement may work against sociometric choice in that the children not praised may feel negatively about those who are.

Within the third grade it will be recalled that sociometric evaluations were obtained for five areas; that is, the children were asked to identify children in the class who they viewed as: (1) most aggressive, (2) most dependent, (3) most achievement oriented, (4) most friendly, and (5) most anxious, although these terms were not used with the children. Thus, the class mean score in each of these areas is taken as an indication of overall class perception along these lines. On the first two dimensions, aggression and dependency, there are only a chance number of relationships with teacher behavior. With the latter three dimensions, however, there are more than the chance number. As shown in Table 4.12, these correlations suggest the following interpretations: First high perception of other members of the class as achievement oriented is associated with teacher behavior which is viewed as confident, intellectually effective, stimulating, non-disparaging, and which provides a

TABLE 4.12

CORRELATIONS BETWEEN CLASS MEAN SOCIOMETRIC SCORES ON MEASURES OF TEACHER BEHAVIOR--GRADE III

	Achievement	Affiliation	Anxiety
Q-Sort Punitive--Total		-.32	
Q-Sort Punitive--Joint		-.36	
Q-Sort Confident--Total	.33		
Q-Sort Confident--Joint	.38		
Q-Sort Intellectual Effectiveness--Total	.37		-.33
Q-Sort Intellectual Effectiveness--Joint	.31		-.49
Q-Sort Disparaging--Total		-.26	
Q-Sort Disparaging--Joint		-.41	
Q-Sort Stimulating--Total			-.30
Q-Sort Stimulating--Joint			-.45
Q-Sort Achievement Oriented--Total			-.37
Q-Sort Achievement Oriented--Joint			-.37
Observations Minimal Reinforcement f	.35		
Observations Minimal Reinforcement %	.29		
Observations Problem Structuring f	.37	.39	
Observations Problem Structuring %	.14	.40	
Observations Ignores Child f		-.31	
Observations Ignores Child %		-.30	
Rating Stimulating	.39	.32	
Rating Disparaging		-.30	
Rating Relaxed	-.33		

good deal of minimal reinforcement and problem structuring; in sum, teacher behavior which while non-disparaging is also well structured and problem oriented. Perception of other members of the class as friendly seems to be associated with teacher behavior which is not punitive or disparaging and which is problem oriented. In this instance, however, as might be expected, the non-punitive variables seem to be of more importance than the intellectual variables whereas the reverse was true with regard to achievement perceptions. Finally, with regard to perceptions of other class members as anxious; this seems to be fostered by teacher behavior which is not intellectually effective nor stimulating nor achievement oriented.

Barron-Welsh

The Barron-Welsh Figures Test, it will be recalled, was administered only in grade three. Only two of the measures of teacher behavior show significant correlations with the Barron-Welsh. However, these seem worth reporting because of their striking agreement, both being measures of non-verbal affiliation. Thus, there are negative correlations with the Q-sort--physical contact--(the correlations both total and joint being -.39)--and with the observation category--'non-verbal affiliation' (the frequency and percentage correlations being -.39 and -.33). On each of these measures, however, these negative correlations also obtained at the beginning of the year though to a lesser degree. Thus, the partial correlation in each case is non-significant. It may well be, however, that this characteristic on the part of the teacher has had some impact upon the pupil behavior as of the time of the fall testing. It is difficult to explain the meaning behind these correlations but the relationship seems quite striking.

Observer Ratings of Class Behavior

It will be recalled that after each observation period the observer recorded not only a rating fot the teacher behavior during that period, but also for the behavior of the class along three dimensions. These were summed across the total of ten observation periods to give a composite rating for the class behavior. Relationships between these ratings and the various measures of teacher behavior are shown in Table 4.13. One cannot, of course, argue that these are independent judgments since the various impressions made by teachers may be in part a function of how the observer felt the class was reacting. This, however, should not be true of the 'observations' category, since the categories here are intended to be objective tallying of teacher behaviors. It is, however, possible for the contrary to be true and for the observer's judgment of the class to be affected by his impressions of teacher behavior. Some evidence that this was not the case, however, is provided by the first rating where it may be noted that the orderliness of the class is not related to the Q-sort as to how controlling the teacher was--though it is correlated with the rating on the 'control' dimension. However, orderliness of the class is correlated much higher with certain other aspects of teacher behavior as assessed; that is, for both grades the extent to which the class was judged as being orderly and work oriented is correlated at a moderately high level--correlations for the most part between .40 and .60 with judgments of the teacher as being confident, achievement oriented, intellectually effective, stimulating, and as exerting less personal control and being more responsive to students seeking attention by raising their hands. In the first grade there is the additional suggestion that the class is more orderly when the teacher

TABLE 4.13

CORRELATIONS BETWEEN MEASURES OF TEACHER BEHAVIOR AND RATINGS OF CLASS BEHAVIOR

	Orderly		Unhappy		Independent	
	Grade I	Grade III	Grade I	Grade III	Grade I	Grade III
Q-Sort Controlling (t)			.66	.72		-.45
Q-Sort Controlling (j)			.58	.68		-.35
Q-Sort Warmth (t)			-.78	-.67		
Q-Sort Warmth (j)			-.64	-.68		
Q-Sort Punitive (t)			.81	.78		-.45
Q-Sort Punitive (j)			.84	.80		-.46
Q-Sort Confidence (t)	.44	.45	-.45			
Q-Sort Confidence (j)	.46	.46				
Q-Sort Supportive (t)			-.81	-.83	.34	
Q-Sort Supportive (j)			-.77	-.78		
Q-Sort Stimulating (t)		.45	-.52			
Q-Sort Stimulating (j)	.31		-.53	-.48		
Q-Sort Achievement Oriented (t)	.52	.49		-.38	.33	
Q-Sort Achievement Oriented (j)	.56	.56				
Q-Sort Physical Contact (t)			-.35	-.40		
Q-Sort Physical Contact (j)						
Q-Sort Intellectual Effectiveness (t)	.52	.60	-.48		.46	
Q-Sort Intellectual Effectiveness (j)	.47	.46	-.30		.34	

TABLE 4.13 (Continued)

	Orderly		Unhappy		Independent	
	Grade I	Grade III	Grade I	Grade III	Grade I	Grade III
Q-Sort Disparaging (t)			.74	.75		-.51
Q-Sort Disparaging (j)			.75	.75		-.43
Q-Sort Factor I (t)			.86	.85		-.39
Q-Sort Factor I (j)			.83	.81		-.34
Observations Answers Hand (f)	.50	.37				.30
Observations Answers Hand (%)	.45	.36				.30
Observations Praise and Encouragement (f)			-.54	-.54		
Observations Praise and Encouragement (%)		-.39	-.45	-.41		.39
Observations Non-verbal Affiliation (f)						-.35
Observations Non-verbal Affiliation (%)					-.33	-.34
Observations Minimal Reinforcement (f)	.60		-.40	-.49		
Observations Minimal Reinforcement (%)	.47		-.40	-.39		
Observations Asking Questions (f)	.63		-.40			
Observations Asking Questions (%)	.44		-.46			
Observations Problem Structuring (f)						
Observations Problem Structuring (%)		.33				

TABLE 4.13 (Continued)

	Orderly		Unhappy		Independent	
	Grade I	Grade III	Grade I	Grade III	Grade I	Grade III
Observations Academic Control (f)			.33			
Observations Academic Control (%)			.41	.32		
Observations Personal Control (f)	-.31	-.43	.81	.55		-.31
Observations Personal Control (%)	-.39	-.47	.82	.56	-.33	-.34
Observations Moralizing (f)					-.34	
Observations Moralizing (%)					-.45	-.33
Observations Hostility and Reprimands (f)		-.31	.86	.69	-.49	-.33
Observations Hostility and Reprimands (%)		-.32	.82	.68	-.45	-.32
Observations Ignores Child (f)			.34		-.34	-.35
Observations Ignores Child (%)			.37		-.34	-.32
Rating Controlling	.38		.43	.64	-.53	-.61
Rating Stimulating	.42	.48	-.81	-.57	.34	
Rating Disparaging			.82	.81	-.43	-.40
Rating Supportive			-.86	-.78		
Rating Confident	.53	.40	-.60			
Rating Affiliative		-.33	-.77	-.69		
Rating Relaxed	-.47	-.53	.52		-.37	

is providing a higher frequency of 'minimal reinforcement' and 'asking question' behavior. These latter two relationships do not hold for the third grade.

The second variable the rating as to the extent to which the class seemed unhappy, shows extremely high relationships with the expected patterns of teacher behavior. In short, those classes were ajudged most unhappy whose teacher was viewed as most punitive, non-supportive, and disparaging--the correlations ranging from .75 to .80. Within the first grade the more objective 'observations' is supportive in that the frequency and percentage of observed hostility and reprimands correlates around .83. This relationship is somewhat lower for grade three, dropping to .69. A similar pattern exists with regard to the category 'personal control' where the correlation is approximately .81 in grade one and .55 in grade three. The observations 'praise and encouragement' category correlates negatively (approximately -.47) in both grades. The correlations with measures of controlling behavior tend to be positive with unhappy judgment of class and the correlations with warmth on the part of behavior negative with the correlations only slightly lower than those for the directly punitive categories ranging around .55 to .70. Thus, there is extremely high agreement between the observer's impressions of a class as being unhappy and perception of the teacher as being punitive, non-supportive, and hostile. Once again, it is important to note that these may not constitute independent judgments.

The third category, 'extent to which the class is judged as able to function in an independent fashion,' shows correlations with teacher behavior at a lower level--ranging from approximately .35 to .55 for the most part and indicates some interesting differences between the two grades. Thus, in the assessment of the observers, independence is negatively correlated with the degree of personal control, moralizing, ignoring the child, and controlling and disparaging behaviors in both grades. In grade one the additional variables which correlate seem to have more to do with the stimulating dimension, suggesting that the stimulating teacher fosters independence whereas in grade three the pattern is more apparent with controlling dimensions, suggesting that the more controlling teacher works contrary to the development of independence. Thus, it appears that excessive hostility and disparaging behavior on the part of the teacher is undesirable at either level if one wishes to foster independence and that in addition to this the stimulating characteristics of the teacher seem to be more important in grade one, whereas the extent of permissiveness seems to be of greater importance in grade three.

Relationships Between Pupil Measures and Teacher Test Measures

Although it is to be expected that most of the relationships with pupil behavior would probably be obtained with the measures of teacher classroom behavior, it is nevertheless of interest to see if there are relationships with teacher test measures.

Achievement Gain

With regard to pupil achievement in grade one,the teacher test measures which relate significantly to one or more measures of pupil achievement are shown in Table 4.14. Of particular interest is the correlation with pupil gain which is determined by obtaining the correlation of the teacher test variable with a post-test achievement measures while partialling out the reading vocabulary pre-test. Those correlations which remain at a significant level after the partial correlation is

TABLE 4.14

RELATIONSHIPS BETWEEN TEACHER TEST MEASURES AND PUPIL ACHIEVEMENT*

	Reading Vocabulary Pre- 117	Reading Vocabulary Post- 118	Reading Comprehension Post- 119	Arithmetic Post- 120
TPS Recognition	-.03	.20 (.213)	.06 (.056)	.30* (.307)
TPS Child's Affection	.07	.25 (.241)	.05 (.053)	.42* (.415)
TPS Child's Autonomy	-.05	.34* (.363)	.14 (.137)	.33* (.338)
TPS Obsessive	.10	.33 (.317)	.13 (.136)	.24 (.230)
TPS Dependency	.16	.28 (.253)	.22 (.238)	.34 (.326)
TPS Exhibitionism	-.04	.16 (.175)	-.03 (-.032)	.33* (.338)
Objective Achievement	-.13	.04 (.074)	.08 (.074)	-.33 (-.318)
Objective Control	-.27	-.00 (.069)	.03 (.018)	-.31 (-.288)
EPPS Achievement	-.32	-.27 (-.175)	-.02 (-.038)	.00 (.044)
EPPS Intraception	-.10	-.46* (-.452)	-.32 (-.327)	-.08 (-.068)
EPPS Nurturance	.32	.12 (.047)	.08 (.101)	.17 (.136)
EPPS Abasement	.22	.44 (.409)	.35 (.371)	.43 (.414)
College	.19	-.36* (-.426)	-.26* (-.255)	-.31* (-.344)
Degree	.08	-.39* (-.423)	-.20 (-.197)	-.25 (-.263)
Major	.38	.18 (.099)	.05 (.075)	-.06 (-.119)
HTP Achievement	-.34	-.27 (-.206)	-.02 (-.039)	.20 (.262)

*Partial r's are in parentheses.

computed are shown by an asterisk. In terms of overall importance to the study it is worth noting that TPS Child Autonomy is generally the best predictor of achievement gain. As has been indicated before, the dimension of teacher behavior, which seems most predictive of achievement gain, is the extent to which the teacher is stimulating in the classroom. This variable was best predicted by the TPS Child's Autonomy score and this score also correlates with achievement gain to provide a coherent picture, the correlations being of the magnitude of .35 with gain in reading vocabulary and arithmetic; with gain in reading comprehension the correlation is considerably lower at .14. Table 4.14 also indicates some other possible teacher test predictors of pupil achievement gain but the correlations do not seem consistent across measures of gain and may simply be artifacts. It is of considerable interest to note that two of the biographical items obtained on the teachers show rather substantial correlations with achievement gain, these being college attended and degree received. In each case these items correlate substantially with all three measures of achievement gain in grade one. With respect to college attended the numerical magnitudes were assigned essentially at random since one cannot argue for a continuous variable applied to different colleges. The nature of the assignment of numbers suggest that within the first grade, the greater gain in achievement is attained by teachers who as students attended the University of Utah. Further, the nature of assigning numbers to the categories for degree possession suggests that the greatest achievement gain is obtained by first grade teachers whose degree holding is limited to the bachelor's degree and that teachers holding a master's degree achieve somewhat less academic gain.

With regard to achievement gain in grade three, there are only a chance number of correlations with tests which are significant. It is of some interest merely to note the one test which does correlate significantly (negatively) with all three measures of achievement gain; and that is the Deference score on the EPPS. This correlation is of the magnitude of -.45 for all three spring achievement measures and is -.22 for the fall achievement measure resulting in partial correlations of -.35 to -.42. The fact that the correlations are of considerable magnitude does suggest that excessive deference on the part of the teacher is conducive to poor achievement gain in grade three; but the fact that this is virtually the only significant correlation to emerge means that it must be treated very tentatively.

Further, the finding of greater gain on the part of University of Utah teachers is repeated for Arithmetic and Reading Comprehension, the partial correlations being .37 and .23 respectively. This result may, of course, be due to differences among these groups of teachers other than the institution attended. It is not, however, attributable to age or years of teaching since these variables show non-significant correlations with achievement gain.

Liking for School

Hypothesis: Measures of Affiliation Need will be negatively correlated with liking--first grade only.

This hypothesis receives some support in that TPS Affiliation correlates -.33 with 'liking for school' and TPS 'Child's Affection' correlates -.38. None of the other affiliation scores correlate significantly, however.

In addition, a number of other measures correlate with liking for school and some of them somewhat higher than the affiliation scales (see Table 4.15). The highest correlation (-.52) is found with TPS Achievement, slightly lower with EPPS Order, and slightly lower with TPS Dominance, Obsessiveness, Status and Control. In total these correlations suggest that pupils in grade one tend to dislike teachers with high achievement, control and order needs as well as high affiliation needs. EPPS Autonomy and Heterosexuality show positive correlations.

TABLE 4.15

'LIKING FOR SCHOOL' AND TEACHER TESTS

	Grade I	Grade III
TPS N. Achievement	-.52**	-.33*
TPS N. Affiliation	-.33*	.08
TPS N. Control	-.27	-.32*
TPS Status Striving	-.38*	- 21
TPS Child's Affection	-.37*	.01
TPS Obsessive	-.39*	-.19
TPS Dominance	-.38*	-.27
Objectives N. Achievement	.02	-.34*
Objectives N. Affiliation	.06	-.32*
EPPS Order	-.47**	-.20
EPPS Autonomy	.31*	-.09
EPPS Heterosexuality	.35*	.12
HTP Recognition	-.43**	-.07
HTP Ego Strength	.36*	.14

*Indicates significance at .05 level.
**Indicates significance at .01 level.

In grade three, four correlations are significant whereas eleven are significant in grade one. It is noteworthy, however, that the pattern of correlations is quite similar for the two grades; the magnitude generally being somewhat higher in the first grade as shown in Table 4.15. In both grades there is a tendency for the teacher to be less liked who is high on measures of achievement and control need and also on the TPS measures of Dominance and Obsessive.

Anxiety

There seems to be little indication that any of the teacher test measures are predictive of extent of pupil anxiety throughout the school year Thus, in grade one there are several tests which correlate significantly with pupil anxiety scores at the beginning of the school year but for the most part these correlations do not hold up at the end of the year. Those scores predictive at the beginning of the year (see Table 4.16) suggest higher anxiety on the part of the classes whose teachers were high in Affiliation, Deference, Endurance, and Order on the EPPS and low on Instrumental Rewards of the TPS scale. There are three measures which show significant correlations at the end of the year with both the test anxiety scale and the measure of general anxiety. These three are the HTP Recognition Need score and the two PSI Affiliation scores. Thus, there may be some suggestion that more anxiety is generated by the teacher having the higher affiliation and recognition needs but the support for this notion is rather weak, particularly since these measures which seem most related to anxiety at the end of the year are tests which seem to predict virtually nothing else by way of pupil or teacher behavior.

In grade three only one of the 32 tests shows a significant correlation with gain in anxiety on both measures, that being the EPPS

TABLE 4.16

RELATIONSHIPS BETWEEN TEACHER TEST VARIABLES AND PUPIL ANXIETY--GRADE ONE

	Anxiety Fall 124	TAS Spring 125	Anxiety Spring 133
TPS Instrumental Rewards	-.33*	-.26	-.07
EPPS Order	.38*	.23	.12
EPPS Endurance	.34*	.36*	.00
EPPS Deference	.36*	.12	.11
EPPS Affiliation	.45*	.04	.23
EPPS Change	-.07	.06	-.32*
EPPS Heterosexuality	-.30	-.31	-.35*
HTP Achievement	-.32	-.03	.02
HTP Recognition	.19	.43	.35*
PSI Affiliation 1	.20	.33	.40*
PSI Affiliation 2	.13	.30	.34*

*Significant at .05 level.

Order score which correlates zero with anxiety scores at the beginning of the year but .34 with the Sarason Test Anxiety scale in the spring and .32 with our questionnaire measure of anxiety.

Ratings of Class Behavior

For the first grade there are only a chance number of relationships between teacher tests and the observer class ratings. For the third grade, however, there are a significant number of correlations which are significant for the ratings on orderliness and unhappiness--not for the rating on independence. These correlations are shown in Table 4.17. These correlations suggest that the class which is rated as an orderly, businesslike class tends to have a teacher with high needs for order, control and

TABLE 4.17

TESTS PREDICTIVE OF OBSERVER RATINGS OF CLASS BEHAVIOR--GRADE III

	Orderly	Unhappy
TPS Affiliation	-.32	
TPS Child's Affection	-.33	
TPS Child's Autonomy	-.34	
TPS Exhibition	-.36	
EPPS Dominance	-.36	
PSI Affiliation	-.35	
EPPS Order	.44	
EPPS Aggression	.35	
HTP Control	.32	
TPS Achievement		.31
TPS Obsessive		.35
Situation Control		.35
EPPS Intraception		-.33
EPPS Deference		-.32
PSI Affiliation		-.42

aggression and low needs for affiliation, child affection, and autonomy or exhibition--a pattern which clearly makes some sense. It will be recalled that the best predictor within grade three of the extent to which the teacher was viewed as stimulating and intellectually effective was the TPS Child's Affection score, the direction being such that a low score on the TPS Child's Affection was considered indicative of a stimulating effective teacher. Thus, one would expect that a low score on this scale would also predict orderliness on the part of the class since this appears to be a function of this teacher characteristic and in fact this is what occurs--this test being one of a cluster which does correlate negatively with orderliness on the part of the class. It is also of interest to note that the best predictor of orderliness on the part of the class is EPPS Order (correlating .44). Although this score is not considered one of the better predictors of teacher behavior, since it has significant correlations with only six criterion measures, it is nevertheless true that five of those six are measures of 'intellectual effectiveness' and 'stimulating' indicating that the teacher high on this score is viewed as intellectually effective and stimulating;and as previously noted, this is the variable which seems to predict an orderly class so that once again the data is consistent on this point.

With regard to judged unhappiness on the part of the class, a consistent pattern once again emerges. It will be recalled that the major variable making for unhappiness as judged by observers is punitiveness on the part of the teacher. Further, punitiveness on the part of the teacher was associated with a high score on TPS Obsessive and low scores on EPPS Intraception and Deference. Thus, one would predict that teachers high on TPS Obsessive or low on EPPS Intraception and Deference would tend to be more punitive and hence have more unhappy classes and these relationships do in fact emerge.[1] The other tests which predict the class ratings also seem to fit within this coherent pattern. It is of interest to note that the three class ratings show quite low correlations among themselves; that is, the correlation between orderliness and unhappiness is -.06, between orderliness and independence -.22, and between unhappiness and independence -.32.

Comparison of Grades (I and III) on Measures of Teacher Behavior and Teacher Tests

With regard to 'observations,' (Table 4.18) the first comparisons of interest pertain to the percentage score. The frequency score cannot be used here since this score was based on distributions within each grade (see page 2.12). This analysis supports the previously discussed finding with regard to between grade differences in minimal reinforcement with significantly more being found in grade three. None of the other major categories show significant differences. However, three categories not assessed through the analysis of variance procedure do show significant 'between grade' differences. First, is the category, 'answering the child's raised hand,' in which the frequency in grade three is nearly twice that in grade one and is highly significant, though its' psychological importance may be questioned. The second category

---

[1] With regard to the class rating on unhappiness, one would expect that this should correlate with the 'liking for school' scale administered to pupils. In the third grade the correlation is significant at -.35 providing some additional validity for both measures. In the first grade, however, the correlation is non-significant and virtually zero casting some question on the validity of these two measures.

TABLE 4.18

COMPARISON OF GRADES I AND III ON MEASURES OF TEACHER BEHAVIOR

	Grade I		Grade III	
	Mean	S.D.	Mean	S.D.
Q-Sort Controlling (Total)[1]	8.22	2.65	7.78	2.70
Q-Sort Controlling (Joint)	4.00	1.55	3.95	1.76
Q-Sort Warmth (Total)	7.16*	3.00	8.72	2.92
Q-Sort Warmth (Joint)	3.58*	1.62	4.35	1.64
Q-Sort Punitive (Total)	8.64*	2.83	7.32	2.82
Q-Sort Punitive (Joint)	4.13	1.73	3.82	1.60
Q-Sort Confident (Total)	8.36	3.10	7.55	2.50
Q-Sort Confident (Joint)	4.08	1.82	3.97	1.56
Q-Sort Supportive (Total)	7.13*	2.87	8.55	2.66
Q-Sort Supportive (Joint)	3.72	1.52	4.25	1.75
Q-Sort Stimulating (Total)	8.27	3.14	7.87	2.62
Q-Sort Stimulating (Joint)	3.97	1.84	4.00	1.54
Q-Sort Achievement Oriented (Total)	8.55	3.24	7.47	2.57
Q-Sort Achievement Oriented (Joint)	4.33	1.84	3.77	1.48
Q-Sort Physical Contact (Total)	6.77**	2.76	9.00	2.47
Q-Sort Physical Contact (Joint)	3.38	1.23	4.50	1.81
Q-Sort Intellectual Effectiveness (Total)	8.66	2.91	7.47	3.00
Q-Sort Intellectual Effectiveness (Joint)	4.50*	1.63	3.60	1.61
Q-Sort Disparaging (Total)	8.25	3.10	7.80	3.05
Q-Sort Disparaging (Joint)	3.86	1.64	4.10	1.69
Observations Answers Hand %	4.30**	3.74	9.00	3.99
Observations Praise and Encouragement %	5.72	2.94	5.08	2.80

TABLE 4.18 (Continued)

	Grade I Mean	Grade I S.D.	Grade III Mean	Grade III S.D.
Observations Non-verbal Affiliation %	.81*	.75	.42	.71
Observations Minimal Reinforcement %	8.33*	2.07	9.57	2.49
Observations Asking Questions %	23.63	5.23	24.52	5.17
Observations Problem Structuring %	29.08	6.14	27.22	6.92
Observations Academic Control %	19.75	4.49	17.12	3.84
Observations Personal Control %	6.38	4.36	5.57	3.64
Observations Moralizing %	1.33*	1.34	.77	.79
Observations Hostility and Reprimands %	.68	1.15	.69	.86
Observations Ignores Child %	.17	.21	.13	.14
Rating Controlling	38.61	7.42	37.67	7.87
Rating Stimulating	36.27	7.41	35.52	6.06
Rating Non-Disparaging	34.11	9.20	32.32	9.01
Rating Non-Supportive	26.94	8.47	29.00	8.46
Rating Confident	34.55	8.09	33.82	7.28
Rating Affiliative	31.61	8.42	28.70	8.88
Rating Relaxed	23.55	7.78	21.80	5.49

* = Significant difference in means at .05 level.

** = Significant difference in means at .01 level.

[1] Due to the Q-Sort procedure, a lower score indicates a higher position on the defining variable.

showing a significant difference is that for 'non-verbal affiliation' in which the frequency in both grades is very low, less than one percent, but significantly higher in grade one; the third, the percentage score for 'moralizing' which again is approximately one percent in both grades but significantly higher in grade one.

With respect to Q-sorts the variables showing a significant difference are the 'warmth' dimension indicating significantly higher rating on this dimension for the first grade teachers; the total Q-sort on 'punitive,' assigning a lower rating to the first grade teachers' a higher score for the first grade teachers on the total 'supportive' Q-sort and a higher score for the third grade teachers on the joint sort of 'intellectually effective.' Finally, the 'physical contact' Q-sort is significantly higher in the first grade on both the total and joint Q-sorts lending further support to the finding of greater physical contact or non-verbal affiliation on the 'observations' measure. None of the 'rating' between grade differences are significant.

Table 4.18 also portrays the percentages of each grade falling in each of the categories on the observation schedule. It is of some interest to note that in both grades teachers were observed to devote approximately 28 per cent of their time to problem structuring, approximately 24 per cent to asking questions, approximately 18 per cent to academic control, approximately 6 per cent to personal control, approximately 9 per cent to minimal reinforcement, approximately 5 per cent to praise and encouragement, around 7 per cent to acknowledging the child's raised hand, about 1 per cent to moralizing, slightly less than one per cent to non-verbal affiliation and to hostility and reprimands.

Turning next to the comparison of the grades with regard to the various teacher test measures, only three of the test measures showed significant differences and since this is almost exactly the number to be expected by chance at the 5 per cent level and since the differences seem to have little meaning, they are not discussed further. With regard to the biographical data on teachers, the significant differences are as follows: The first grade teachers are on the average approximately 10 years older than the third grade teachers, averaging 51 years of age whereas the third grade teachers average 41 years of age. This very likely accounts for the significant differences on two other measures indicating that the first grade teachers have significantly more years of teaching experience (approximately 9 more years) and significantly more academic hours past graduation.

CHAPTER V

TEACHER-PUPIL INTERACTION

Analysis of Variance

The analysis of variance procedure was utilized for the purpose of testing interactions between pupil characteristics and teacher characteristics. The teacher characteristic utilized was that of controlling behavior as assessed by 'observation' scores. To select teachers' scores, in categories six and seven, 'academic control' and 'personal control' respectively, were combined. Within each of the two grades the ten highest and ten lowest teachers on this score were selected as the high controlling and low controlling teachers, each group consisting of slightly less than one-third of the total number of teachers at each grade level. After the ten high and low teachers had been selected within each grade, they were subdivided so as to give additional subgroupings of five teachers each. The purpose of this was to replicate the analyses. Thus, for each analysis we were able to compare five high control teachers with five low control teachers and this design is replicated with a second set of teachers.

Within each of these classes the sexes were differentiated and analyses performed separately for each sex. The two student characteristics utilized were the pre-test or fall scores on (1) reading vocabulary, (2) dependency questionnaire. In each case all students of the same sex within each class were grouped and the scores scrutinized so as to select out initially the highest and lowest five of each sex on the particular variable of concern. As was to be expected, there was considerable overlap from class to class as to the scores distinguishing the high and low groups. Thus, in several instances one of the students in the low group in a particular class would have a score considerably higher than some of the scores in the high group of another class. In order that the distinction, high and low, as applied for example to reading vocabulary have consistent meaning across all classes as opposed to simply within each class, cutting scores were established as so to minimize the overlap.

Grade Three

For the third grade, reading vocabulary scores, the number of each sex per classification per class was reduced to four, i.e., within each class four boys were selected as the high group on reading vocabulary in the fall, etc. For the girls a cutting score at 17 was utilized such that scores at 17 and under were placed in the low group, scores of 18 and over in the high group. This score effectively separated the two groups, the only overlap being two cases in the low group with scores of 18 and two cases placed in the high group, one with a score of 16 and one with a score of 14. The only other criterion used in selecting cases within each class was that extremely atypical scores were avoided. Thus, scores of 0, indicating no correct answers, were not utilized where possible nor were scores which were extremely high and unusually so. For the third grade boys a cutting score at 13 was used, i.e., scores of 13 and below were placed in the low group, scores of 14 and above in the high group. This procedure resulted in only two exceptions, one student placed in the low group with a score of 14 and one student placed in the high group with a score of 13. The range of scores on this measure was from 0 to 42; the variability within each pupil--sex group--within each class was roughly comparable from class to class though not precisely so. As can be seen from the cutting scores, there tends to be greater variability among the "high" group of each sex than among the "low" group, though this is more true for boys.

With regard to the selection of cases based on the fall dependency scale the same basic procedure was followed, i.e., an attempt was made to establish cutting scores which would eliminate overlap among scores in different classes. It was found that in order to provide decent discrimination between high and low scores for girls, it was necessary to reduce the group size to three. Since, however, there are two groups of teachers involved in each analysis, it was possible to divide the teachers in such a way that for one analysis an N of three per group per class could be utilized whereas for the second analysis an N of four could be utilized. In both cases a cutting score of 7 was used, i.e., scores of 7 and below on the dependency questionnaire were placed in the low group and score of 8 and above were placed in the high group. This procedure resulted in no cases which were exceptions to this cutting score. With respect to the boys' scores two groups of teachers again were utilized as with the girls but in this case the N per group per class for one group was four whereas for the other group it was five. For the boys a cutting score of 6 was used such that scores of 6 and below were placed in the low group and above 6 were placed in the high group. This procedure resulted in no exceptions and no overlap of high and low groups for the analysis utilizing five cases per cell. For the analysis using 4 cases per cell there are four cases which are exceptions, i.e., two cases with scores of 7 which are placed in the low group and two scores of 6 placed in the high group.

To summarize, each analysis consisted of a comparison of teacher characteristics as one main effect with a group (4 or 5) of high controlling teachers being compared to a same size group of low controlling teachers. A second main effect was pupil characteristic. In one set of analyses reading vocabulary was used as the selection variable and in the second set dependency questionnaire scores were used as the selection variable. Analyses were done separately for each sex with the procedures described above resulting in equal N's per cell which varied from two to five depending upon the analysis. The above procedure was then replicated with a second set of teachers. In each case the analysis of variance table is identical and is presented in Figure 5.01. One of the teachers in the high group had to be eliminated since the distribution of scores was much higher than for the other classes. Accordingly one class from the low group which also had a high distribution of scores was deleted; and these classes were deleted from all analyses. One other consideration governed the selection of students in the various classes. Since there was some question as to whether the vocabulary test would allow sufficient ceiling in the spring testing and since gain on this measure was one of the dependent variables of particular interest, extremely high scores on the fall testing on reading vocabulary were deleted where possible. In general scores above 37 on the fall testing were eliminzted. There were two exceptions to this where scores higher than this could not be deleted and maintain the other criteria.

Grade One.--The pattern for the grade one analyses was the same as for grade three. However, the smaller number of cases per grade as well as the distribution on the achievement measure made the selection of cases more difficult. Once again the total group of teachers was divided into two groups of high and low--one group containing five high control teachers and five low control teachers; the second group containing four of each. For one of these groups of teachers, here called the "A" group, the number of children per sex per pupil measure was two throughout; that is, for each of the teachers in the "A" group there were two boys classified in the high group on reading vocabulary and two boys (not necessarily the same two) classified in the high group on dependency and similarly for the other three cells. For the "B" group of teachers, consisting of four each in the high and low control groups, the number of cases per cell was three throughout. On reading vocabulary for the girls a cutting score of 23 was utilized resulting in two misclassified cases in the "A" group; for the "B" group one case was misclassified. For the boys a cutting score of 22 was used resulting in 4 misclassified cases in the "A" group. On the dependency measure a cutting score of 6 was used for boys; that is, scores of 6 and below were placed in the low dependency group; for the "B" group of teachers this resulted in four misplacements; for the "A" group of teachers no misplacements. For the girls a cutting score of 7 and below placed pupils in the low dependency category; and this score resulted in a misplacement of two students in the "B" teacher group and four students in the "A" teacher group. The classes of two teachers who were originally included when the top and bottom ten teachers on control were selected were deleted because of extremely atypical distributions which precluded utilizing these or other cutting scores and placing students with any consistency.

FIGURE 5.01

ANOVA DESIGN

			High Control Teachers N = 4 or 5					Low Control Teachers N = 4 or 5				
			Tch 1	Tch 2	Tch 3	Tch 4	Tch 5	Tch 6	Tch 7	Tch 8	Tch 9	Tch 10
High Ability Pupils	or	High Dependency Pupils	N = 2 to 5									
Low Ability Pupils		Low Dependency Pupils										

This design permits the following ANOVA:

	d.f.
Between Teachers	9 (assuming 10 teachers)
Between High and Low Control Teachers	1
Between Teachers Within Control Groups	8
Between Pupil Groups	1
Interactions: Teacher X Pupil Group	9
Control X Pupil Group	1
Residual Among Cells	8
Within Cells	N - 19
Total	N - 1

Due to problems encountered in utilizing this analysis with computer programs--due to missing data--the design was collapsed to the following and cell means were substituted for missing scores.

		High Control Teachers	Low Control Teachers
High Ability Pupils	High Dependency Pupils	N = 10 to 25	
or			
Low Ability Pupils	Low Dependency Pupils		

ANOVA:

	d.f.
Between 'Control' Groups	1
Between 'Pupil' Groups	1
Interaction	1
Within Cells	N - 4
Total	N - 1

It should be noted that these analyses do not involve random assignment of pupils to teachers (or teacher-type) raising some question as to the legitimacy of the analyses. Although assignment of pupils to teachers in practice can hardly be considered to be random (due to geographical location of schools, etc.), one can argue that assignment of pupils to teachers within schools (at least in this study) is independent of the variables we have used to classify pupils; thus lending some justification to treating the data 'as if' random assignment had occurred.

Once the teachers and pupils had been selected, the ANOVA was performed using the dependent variables listed below. Regressed gain scores were used where the pre-/post- correlation warranted; in the other instances post-test scores were used.

GRADE I

1. Gain--Reading Vocabulary
2. Gain--Circles Test
3. Post-test (Spring)--Reading Comprehension
4. Post-test (Spring)--Arithmetic Fundamentals
5. Post-test (Spring)--Sarason Test Anxiety
6. Post-test (Spring)--Questionnaire--Perceived Affiliation (of teacher)
7. Post-test (Spring)--Questionnaire--Perceived Control (of teacher)
8. Post-test (Spring)--Questionnaire--Perceived Achievement (of teacher)
9. Post-test (Spring)--Questionnaire--Liking for School
10. Post-test (Spring)--Questionnaire--Anxiety Scale
11. Post-test (Spring)--Circles Test--Total Score
12. Post-test (Spring--Sociometric Ratings Received

GRADE III

1. Gain--Reading Vocabulary
2. Gain--Reading Comprehension
3. Gain--Arithmetic Fundamentals
4. Gain--Questionnaire Anxiety Scale
5. Gain--Barron-Welsh
6. Gain--Circles Test--Rating
7. Post-test (Spring)--Sarason Test Anxiety
8. Post-test (Spring)--Questionnaire--Perceived Affiliation (of teacher)
9. Post-test (Spring)--Questionnaire--Perceived Control (of teacher)

10. Post-test (Spring)--Questionnaire--Perceived Achievement (of teacher)

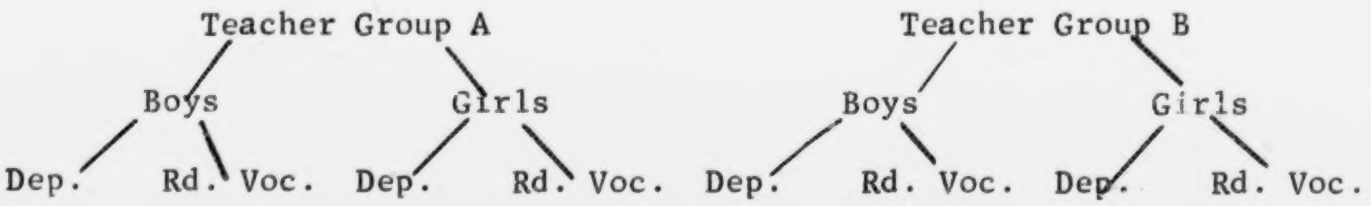
11. Post-test (Spring)--Questionnaire--Liking for School

Results

Based on our prior research, one hypothesis was put forth as follows: H: Within grade *one* pupils scoring high on the questionnaire measure of 'dependency' will show greater gain in Reading Vocabulary under high controlling teachers whereas children scoring low in 'dependency' will show greater gain with the low controlling teachers.

This hypothses received no support from the data. Interaction F's were non-significant for both boys and girls in each replication (A and B groups of teachers and pupils).

It will be recalled that each dependent variable was studied in eight analyses per grade as shown in the following diagram:



Thus, for grade one, ninety-six analyses were performed and for grade three eighty-eight. It will further be recalled that it was the teacher-pupil interaction term which was of interest. One would expect four to five significant interaction F's by chance at the .05 level; in-fact, eight were obtained in grade one and seven in grade three, numbers which are within the .05 limits of chance departure from an expected value of 5. Since, however, the A and B groups constitute replications, one may have considerable confidence that an interaction which is significant (.05 level) in *both* replications represents more than chance. In our results, this occurred only once and hence is the only finding in which we have confidence. As shown in Tables 5.01 through 5.04, these results suggest that in grade three high ability girls (Reading Vocabulary Pre-Test) improve their reading vocabularies more under low controlling teachers whereas low ability girls improve more with high controlling teachers.

TABLE 5.01

ANOVA--GRADE THREE GIRLS SELECTED ON READING VOCABULARY--
GROUP A; DEPENDENT VARIABLES IS IN
READING VOCABULARY

Source	S.S.	d.f.	Mean Sq.	F
Between High and Low Control Teachers	3.16	1	3.16	
Between High and Low Vocabulary Pupils	25.42	1	25.43	
Interaction I.J.	97.90	1	97.90	4.09*
Error	1817.80	76	23.92	

*Significant .05 level.

TABLE 5.02

ANOVA--GRADE THREE GIRLS SELECTED ON READING VOCABULARY--GROUP B; DEPENDENT VARIABLE IS GAIN IN READING VOCABULARY

Source	S.S.	d.f.	Mean Sq.	F
Between High and Low Control Teachers	20.76	1	20.76	
Between High and Low Vocabulary Pupils	47.90	1	47.90	
Interaction I.J.	142.31	1	142.31	4.17*
Error	2593.14	76	34.12	

*Significant .05 level.

TABLE 5.03

MEANS--GRADE THREE GIRLS SELECTED ON READING VOCABULARY PRE-TEST--GROUP A; DEPENDENT VARIABLE IS GAIN IN READING VOCABULARY

	High Control Teachers		Low Control Teachers	
	Mean	S.D.	Mean	S.D.
High Vocabulary	- .92*	4.56	1.87	5.12
	(-1.15)**		(1.17)	
Low Vocabulary	2.52	3.92	.68	6.23
	(1.15)		(-1.17)	

*These are derived values having no intrinsic meaning.

**Discrepancies from expected cell means based on marginal and grand means.

TABLE 5.04

MEANS--GRADE THREE GIRLS SELECTED ON READING VOCABULARY PRE-TEST--GROUP B; DEPENDENT VARIABLE IS GAIN IN READING VOCABULARY

	High Control Teachers		Low Control Teachers	
	Mean	S.D.	Mean	S.D.
High Vocabulary	-2.78*	7.93	1.03	3.60
	(-1.39)**		(1.38)	
Low Vocabulary	1.53	5.25	- .20	5.97
	(1.39)		(-1.38)	

*These are derived values having no intrinsic meaning.

**Discrepancies from expected cell means based on marginal and grade means.

CHAPTER VI

SUMMARY AND DISCUSSION

A sample of teachers in each of grades one (N = 36) and three (N = 40) was obtained within a large metropolitan school district. Relationships between teacher measures and pupil (class) change and status measures were of primary interest. Several hypotheses based directly on our prior research were tested. In addition, consistencies among measures of teacher behavior and relationships between teacher behavior and teacher test data were studied. Finally, the interactions of teacher 'control' and the pupil characteristics of 'dependency' and 'academic ability' (Reading Vocabulary) were studied.

Teacher measures included (a) categorizations of classroom behavior during each of nine observation period, (b) Q-Sorts and Ratings of observer judgments following observation, and (c) test scores on several questionnaire and semi-projective devices. Pupil measures included regressed gain scores on the California Achievement Test (Reading Vocabulary, Reading Comprehension, and Arithmetic Fundamentals); Torrance 'Circles' Test; a questionnaire (which was read to the pupils) including the Sarason Test Anxiety Scale, the Medley and Klein 'Liking for School' scale and several other scales developed on the project. A sociometric device was administered in the Spring only. Observations of class behavior were also obtained. Analysis consisted of first order and partial correlations and factorial analysis of variance.

Adequacy of Observational Measures

The evidence seems to indicate that we were quite successful in achieving measures of teacher behavior which provided meaningful differentiation among teachers. The most objective of these, the 'observations' of classroom behavior in which tallies were made within specified categories showed very adequate agreement among observers when observing the same teacher at the same period of time and demonstrated that when different observers observed the teacher over a period of eight observation sessions meaningful differences among teachers which accounted for substantial proportions of the total variance of the observations did emerge. With regard to the correspondence between the observations and the Q-sorts and ratings the degree of agreement on corresponding variables was gratifying as was that between the ratings and Q-sorts and rating variables--and the 'observations' categories primarily pertaining to the Q-sort variables of warmth, supportiveness, and punitiveness. There are, however, no clear *a priori* relationships which might have been specified between the observations and the Q-sort variables of stimulating, achievement oriented, and intellectually effective Since these characteristics of teachers were found to be of considerable significance with relation to certain pupil measures, it is appropriate to note that there is within both grades some degree of correspondence between the 'stimulating' Q-sort and the observations of 'praise and encouragement' and 'asking questions' (positive correlations) and 'personal control,' and 'hostility and reprimands' (negative correlations). Also, the Q-sort 'intellectually effective,' is negatively correlated with 'personal control' observations in both grades. 'Achievement orientation' in the third grade is essentially unrelated to any of the observation categories and in grade one relates only to the asking questions category. It is likely that these aspects of teacher behavior: 'stimulating,' 'intellectually effective,' and 'achievement oriented' are somewhat more difficult to assess through straightforward tallying of described behaviors.

Prediction of Teacher Behavior From Tests

Although there is mild support in our data for the hypothesis that the TPS control score would predict the dimension of controlling behavior and to a lesser extent affiliative behavior on the part of the teacher the confirmatory evidence is quite weak and disappointing in that these relationships had been so clear-cut in two previous samples. In our present data it is clear that certain tests are predicting aspects of teacher behavior quite effectively though one cannot say how well these would hold up in subsequent studies. The only test which functions in a similar fashion across the two grades is the EPPS dominance scale which in both grades predicts the 'warmth' and 'controlling' dimensions with some adequacy, the correlations ranging from .25 to .45 in magnitude. The direction is such that in both grades the teacher scoring high on EPPS dominance is predicted to exhibit more warmth and supportive behavior and less controlling punitive-type behaviors. From a theoretical standpoint there seems no clear reason to expect this type of behavior on the part of the teacher with high dominance needs. Several other tests within each of the two grades seem to predict this general domain of behavior about as well. Within grade one these tests are TPS child autonomy, EPPS autonomy, EPPS aggression, EPPS deference and PSI control. Overall the best predictor within the first grade of the controlling and warmth dimension is the PSI control score though it does not function well in grade three. Other scores within grade three which predict the 'warmth-controlling' character of teacher behavior are the EPPS nurturance, deference, and intraception scales. The EPPS deference scale was also a predictor in grade one but tends to function in the opposite direction in grade three. The best predictor in grade three of these teacher behaviors is the 'situations control' score though this does not function well in grade one. The suggestion is made that a very situationally oriented, obvious type of test is the best predictor in the upper grades such as grade three;and this finding is consistent with our previous research which found that a straightforward type of questionnaire seemed to be the best predictor. Within grade one we suggest that a more subtle projective device operates more effectively.

A teacher behavior characteristic of considerable importance is that dealing with the 'stimulating,' 'intellectually effective,' and 'achievement oriented' domain. Within the first grade these measures are predicted fairly adequately, with correlations of magnitude .25 to .45, by the TPS child's autonomy scale though once again there is no clear theoretical reason for this. The direction of correlations indicates that the teacher scoring higher on child's autonomy is viewed as more stimulating and more intellectually effective. In grade three the best predictor of this dimension appears to be the TPS child's affection score; the direction of correlations indicating that teachers scoring high on this scale tend to be viewed as less stimulating and achievement oriented. In summary then it is our feeling that the tests utilized in this study do provide some basis for predicting teacher behavior within the first and third grades. It is somewhat surprising, however, to find that the nature of the tests which predict is quite different for the most part in the two grades and finally somewhat distressing to note that most of the predictors make little theoretical sense.

Relationships Between Teacher Behavior and Pupil Behavior

With regard to relationships between teacher behavior and pupil characteristics eight hypotheses were developed based on our prior research of which four may be stated to have been clearly supported by present data. Two received some support and two received no support at all. In order these are as follows:

1. Positive correlations between achievement gain and extent to which the teacher was viewed as stimulating; both grades one and three--clearly supported.
5. Positive correlations between liking for school and extent to which the teacher is viewed as warm and permissive; grade three only--clearly supported.
3. Positive correlations between overt affiliative behavior on the part of the teacher and gain in anxiety; grade one--clearly supported.
7. Liking for school will be negatively correlated with measures of affiliation; first grade only. This hypothesis was supported when applied to the index of physical contact or nonverbal affiliation. Although the correlation was not high, it was significant in the predicted direction .29 and supports

the hypothesis that this type of affectionate behavior on the part of the first grade teacher results in her being less well liked by the pupils. The other measures of general affiliation and warmth did not support this hypothesis.

2. Negative correlations between supportive behavior on the part of the teacher and increase in anxiety; grades one and three--some support, grade three only.
8. Negative correlations between extent of controlling behavior on the part of the teacher and gain on the Torrance Circles Tests--some support.
4. Liking for school negatively correlated with degree of achievement orientation as assessed by observers; grades one and three--not supported.
6. Liking for school positively correlated with extent of controlling behavior on the part of the teacher; grade one--not supported.

In addition to these hypotheses the following empirical findings emerged within each of the pupil areas:

Achievement Gain.--In both grades gain in achievement was related to observer impressions of intellectual effectiveness and achievement orinetation on the part of the teacher as well as extent to which she was viewed as stimulating. As has been noted, the observations measure does not directly relate to these Q-sort variables. However, the 'asking questions' category does show significant correlations with gain in reading in the first grade and arithmetic gain in the third grade. There is also the suggestion that extent of personal control exerted by the teacher has a negative correlation with gains in the first grade and that the extent of teacher confidence is positively related with gains in grade one.

Gain in Anxiety.--There are few additional correlations of anxiety gain with teacher behavior. It is of interest to note, however, that in the first grade there are several significant correlations with teacher behavior at the time of the pre-test which may indicate that the teacher characteristics have had their impact in the first six weeks of school. These correlations suggest that 'achievement orientation' and 'confidence' on the part of the teacher is negatively correlated with anxiety. Finally, the one correlation which is significant with anxiety gain is that of praise and encouragement in the first grade. This correlation makes little theoretical sense and may be a chance finding.

Liking for School.--No additional findings.

Change on the Circles Test of Divergent Thinking.--In grade three there is a strong indication that gain on the Circles Test is fostered by the teacher who is intellectually effective and achievement oriented.

Peer Evaluation.--Within the first grade a higher level of generalized liking in the class seems to be related to teacher behavior which is viewed as intellectually effective and prone to moralizing which we interpret to the effect that the moralizing done by the first grade teacher as to how one has to get along socially is probably quite effective.

Within the third grade, sociometric choices were made within five areas, two of which, aggression, and dependency, show only chance relationships with teacher behavior. Perception of other members of the class as achievement oriented is associated with teacher behavior which while not disparaging, is well structured and problem oriented. Perception of other members of the class as friendly is associated primarily with teacher behavior which is viewed as non-punitive. Perception of other members of the class as anxious seems to be fostered by teachers who are viewed as 'intellectually ineffective,' 'unstimulating' and not 'achievement oriented.'

Observation of Class (Pupils).--After each observation period, the class was rated on three dimensions--orderliness, unhappiness, and independence. Summations of these ratings were studied in relation to teacher behavior. Within both grades orderliness was related to the teacher characteristics of 'stimulating,' 'intellectual effectiveness' and 'achievement oriented' whereas unhappiness was related to the teacher characteristics of 'punitive,' disparaging' and 'non-supportive' with correlations of the magnitude .60 to .80. Independence was negatively correlated with 'personal control,' 'moralizing,' 'control' and 'disparaging' in both grades but at a somewhat lower level (.35 to .55). Further, independence in grade one was related to the teacher characteristics 'stimulating' and 'intellectual effectiveness.' It must be kept in mind that these various observations are not independent in that they are all based on observation of teacher-pupil interaction and, with the exception of the presumably objective categorization of teacher behavior, reflect observer impressions.

Teacher Tests and Pupil Behavior

With regard to achievement gain the principle finding is that in grade one the TPS child's autonomy score which was found to predict the degree to which the teacher was stimulating also predicts extent of achievement gain, thus providing a coherent picture in that the teacher scoring high on this scale was likely to be viewed as more stimulating and the more stimulating teacher achieved a greater achievement gain. With regard to 'liking for school' the pattern is that teachers scoring high on a number of TPS needs are liked less well by the pupils, particularly in grade one; the highest correlation being -.52 with TPS need achievement. There seem to be no substantial correlations between the teacher test scores and pupil change in anxiety.

Ratings on three dimensions of class (pupil) behavior show only chance correlations with teacher tests in grade one. In grade three, however, the 'TPS child's affection' score (which correlates with the 'stimulating' dimension of teacher behavior which in turn correlates with 'orderliness' of class behavior) correlates negatively as expected, with 'orderliness.' Also, three scales (TPS Obsessive, EPPS Intraception, and EPPS Deference) which correlate with 'punitiveness' of teacher behavior correlate as would be expected with 'unhappiness' of the class as judged by observers.

In summary, if one views gain in achievement, liking for school, less anxiety, and increase in 'divergent thinking' as desirable school outcomes, it appears that in the third grade one should obtain teachers who are viewed as stimulating and intellectually effective while at the same time being viewed as warm, supportive persons; these not being incompatible characteristics. In the first grade it appears that one should once again attempt to obtain stimulating, intellectually effective teachers while at the same time guarding against teachers who are overtly demonstrative in their affection for students and also those who may have extremely strong personality needs particularly in the areas of achievement need and control need. Further, it appears that certain psychological tests offer considerable promise in the prediction of such teacher behavior. Although most of the correlations reported in this study are of modest magnitude, ranging for the most part from .32 to .55 with the majority being under .40, it is felt that, considering the difficulties in obtaining adequate measures of teacher behavior combined with the difficulties in assessing change on the part of pupils, these relationships are of considerable importance.

Teacher-Pupil Interaction

Analysis of Variance procedures were used to assess the effect upon our various dependent variables of pupil behavior of the interaction of the teacher characteristic--'controlling' with two pupil characteristics: academic ability (as measured by pre-test scores on Reading Vocabulary) and dependency (as measured by a questionnaire).

One hypothesis based on our prior research, i.e., more dependent children would show greater achievement gain with the more controlling teacher--first grade only--was not supported The major finding of these analyses, in which we have considerable confidence since it was replicated across two groups of teachers and pupils is that among third grade girls greater gain in Reading Vocabulary occurred among high ability girls with low 'controlling' teachers and among low ability girls with high controlling teachers.

APPENDIX A

PUPIL QUESTIONNAIRE (Keyed)

1. Do you worry when the teacher says that she is going to ask you questions to find out how much you know? y*

2. Do you worry about being promoted; that is, passing from the _____ to the _____ grade at the end of the year? y

3. When the teacher asks you to get up in front of the class and read aloud, are you afraid that you are going to make some bad mistakes? y

4. When the teacher says that she is going to call upon some boys and girls in the class to do arithmetic problems, do you hope that she will call upon someone else and not on you? y

5. Do you sometimes dream at night that you are in school and cannot answer the teacher's questions? y

6. When the teacher says that she is going to find out how much you have learned, does your heart begin to beat faster? y

7. When the teacher is teaching you about arithmetic, do you feel that other children in the class understand her better than you? y

8. When you are in bed at night, do you sometimes worry about how you are going to do in class the next day? y

9. When the teacher asks you to write on the blackboard in front of the class, does the hand you write with sometimes shake a little? y

10. When the teacher is teaching you about reading, do you feel that other children in class understand her better than you? y

11. Do you think you worry more about school than other children? y

12. When you are at home and you are thinking about your arithmetic lesson for the next day, do you become afraid that you will get the answers wrong when the teacher calls upon you? y

13. If you are sick and miss school, do you worry that you will do more poorly in your schoolwork than other children when you return to school? y

---

*Indicates answer which adds to score.

14. Do you sometimes dream at night that other boys and girls in your class can do things you cannot do? y

15. When you are home and you are thinking about your reading lesson for the next day, do you worry that you will do poorly on the lesson? y

16. When the teacher says that she is going to find out how much you have learned, do you get a funny feeling in your stomach? y

17. If you did very poorly when the teacher called on you, would you probably feel like crying even though you would try not to cry? y

18. Do you sometimes dream at night that the teacher is angry because you do not know your lessons? y

In the following questions the word "test" is used. What I mean by "test" is any time the teacher asks you to do something to find out how much you know or how much you have learned; it could be by your writing on paper, or by your speaking aloud, or by your writing on the blackboard. Do you understand what I mean by "test"--it is any time the teacher asks you to do something to find out how much you know.

19. Are you afraid of school tests? y

20. Do you worry a lot before you take a test? y

21. Do you worry a lot while you are taking a test? y

22. After you have taken a test do you worry about how well you did on the test? y

23. Do you sometimes dream at night that you did poorly on a test you had in school that day? y

24. When you are taking a test, does the hand you write with shake a little? y

25. When the teacher says that she is going to give the class a test, do you become afraid that you will do poorly? y

26. When you are taking a hard test, do you forget some things you know very well before you started taking the test? y

27. Do you wish a lot of times that you didn't worry so much about tests? y

28. When the teacher says that she is going to give the class a test, do you get a nervous or funny feeling? y

29. While you are taking a test do you usually think you are going to do poorly? y

30. While you are on your way to school, do you sometimes worry that the teacher may give the class a test? y

ND 31. Do you like to tell your troubles to your teacher? y

PA 32. Do you think of your teacher as a friend? y

PC 33. Do you have to do lots of things in school that you don't want to do? y

PC 34. In school are you always told what to do and when to do it? y

PAch 35. Does your teacher expect everyone to do their very best? y

PC 36. Do you help plan what the class is going to do? n

PA 37. Would you like to do with your teacher outside of school? y

PA 38. Is your teacher ever mean? n

PAch 39. Do you always find out whether your work is right? y

PC 40. Can you leave your seat without asking? n

PA 41. Do you hope all your teachers are like the one you have now? y

PC 42. Do children ask a lot of questions in class? n

PAch 43. Is your teacher mostly interested in how much you learn? y

PAch 44. Is your teacher very often wrong? n

PA 45. Is your teacher mostly interested in whether you are happy? y

PAch 46. Do you think you have learned a lot from your teacher? y

L 47. Do you ever feel like staying away from school? n

L 48. Do you have much fun in this class? y

L 49. Do you always do your best in this class? y

L 50. Do most of the children like the teacher? y

L 51. Does the teacher help enough? y

ND 52. Do you like to be told exactly how to do things? y

NA 53. Does it upset you to think that you are not liked by everyone? y

NAch	54.	Do you like to look at books outside of school?	y
NAch	55.	Do you like to work with numbers?	y
NA	56.	Would you give up your recess to help a friend?	y
Anx	57.	Are you afraid of going to new places?	y
NA	58.	Do you like to be around lots of children?	y
ND	59.	Do you like to have to figure things out by yourself?	n
NA	60.	Do you care what other people think of you?	y
NAch	61.	Do you like to find out new things by yourself?	y
Anx	62.	Do lots of things frighten you?	y
NA	63.	Do you like to work alond?	n
NAch	64.	Do you like to learn new things even if it's hard?	y
NAch	65.	Would you give up your recess to get a better mark?	y
	66.	Are you afraid of animals?	y
ND	67.	Do you agree with everything your teacher says?	y
NA	68.	Does it bother you if other children have more friends than you do?	y
ND	69.	Do you think children should always do what grown-ups tell them to do?	y
Anx	70.	Are you afraid to do things because you might get hurt?	y
ND	71.	Do you think children should always agree with their parents?	y
NA	72.	Are you unhappy if you have to walk home from school alone?	y
NAch	73.	Do you like to take tests to show how much you know?	y
Anx	74.	Does it frighten you to be left alone?	y
ND	75.	Do you like to be the leader when you play games?	n
ND	76.	Do you like to feel that your teacher will always take care of you?	y
NAch	77.	Are you usually happy with your school work?	n

NAch	78.	Do you like to ask questions?	y
NA	79.	Does it make you feel bad to get angry at someone else?	y
Anx	80.	Does it bother you to go to a party where you don't know many of the children?	y
ND	81.	Does it sometimes bother you when grown-ups want to help you do things?	n
NAch	82.	Do you like to do extra work?	y
Anx	83.	Does it make you unhappy when someone tells you did something wrong?	y
NAch	84.	Do you often wonder why things happen the way they do?	y
ND	85.	Do you like to have your teacher help you with your work?	y
NA	86.	Do you like a lot of people to call you dear?	y
Anx	87.	Were you afraid on the first day of school?	y
Anx	88.	Does it upset you to hear people argue?	y
NA	89.	Do you like to do things with your classmates rather than by yourself?	y
ND	90.	Do you like to tell your troubles to your teacher?	y

TEACHER BEHAVIOR

1. Acknowledges students raised hand.--When teacher calls sequences of names (acknowledging hands) record about half of actual acts.

2. Praise and encouragement.--Supportive behavior. Positive reinforcement. Value judgements included. Examples: Very good, that's right, find.

2a. Non-verbal affiliation.--Physical contact with student, such as putting arm around student.

3. Minimal reinforcement.--Examples: Uh huh, okay, all right. Smiles at student. Acknowledge about half when in rapid sequences (as in 1).

4. Asking questions with intent that the student answer.--Example: How many pennies make a dollar? (If teacher calls on student with raised hand, "1" is also scored. If teacher calls on student without his hand being raised, a "6" is scored.)

5. Explaning or problem structuring.--Helps with words while reading, clarifying material; directive statements closely tied to content of material being taught.

6. Academic control.--Teacher directs students to perform certain actions clearly related to academic learning. Includes rhetorical questions. Examples: Open your books; come to my desk for help; read page 13. Calls on student who doesn't have his hand raised.

7. Personal control.--Teacher directs students to perform or stop certain actions related to personal behavior. Examples: Teacher rearranged pupil's chair; lay your pencil down when you are through;

put your hands on your head when you are finished; go to your seat; sit up straight. Facial expression included.

7a. Moralizing by teacher.--Example: Don't do that, you wouldn't like it if Johnny did that to you.

8. Hostility and reprimands.--Example: Shut up! Strikes child.

8x. Ignores child.--

APPENDIX B

FIRST GRADE TEACHER MATRIX

Card No.	Variable Number	Name of Variable
Identification		
1		Grade
1		School
1		Teacher No.
Teaching Preference Schedule		
1	1	N - Achievement
1	2	N - Affiliation
1	3	N - Recognition
1	4	N - Control (Constant = +40)
1	5	GS1 - Instrumental Rewards
1	6	GS2 - Status Striving
1	7	GS3 - Child's Affection
1	8	GS4 - Child's Autonomy
1	9	GS5 - Rebellious Motives
1	10	GS6 - Vicarious Motives
1	11	GS7 - Obsessive
1	12	GS8 - Dependency
1	13	GS9 - Exhibitionism
1	14	GS10 - Dominance
Personal Preferences for Educational Objectives		
1	15	N - Achievement
1	16	N - Affiliation
1	17	N - Recognition

Card No.	Variable Number	Name of Variable
1	18	N - Control
Edwards Personal Preference Schedule		
1	19	Achievement
1	20	Order
1	21	Autonomy
1	22	Intraception
1	23	Dominance
1	24	Nurturance
1	25	Endurance
1	26	Aggression
1	27	Deference
1	28	Exhibition
1	29	Affiliation
1	30	Succorance
1	31	Abasement
1	32	Change
1	33	Heterosexuality
1	34	Cons
Identification		
2		Grade
2		School
2		Teacher No.
Situations Test		
2	35	N - Achievement
2	36	N - Affiliation
2	37	N - Recognition
2	38	N - Control

Card No.	Variable Number	Name of Variable
Identification		
Biographical Data		
2	39	Age
2	40	Years Taught
2	41	College
2	42	Degree
2	43	Hours Over Graduation
2	44	Year Degree Received
2	45	Major
Problem Situation Inventory		
2	46	Control
House - Tree - Person		
2	47	Control
2	48	Affiliation
2	49	Achievement
2	50	Recognition
2	51	Ego
Problem Situation Inventory		
2	52	Item (Affiliation)
2	53	Q (Affiliation)
Identification		
2	Card No.	
Identification		
3		Grade
3		School

Card No.	Variable Number	Name of Variable
Identification		
3		Teacher No.
Q-Sorts		
3	54	Controlling - Total
3	55	Controlling - Joint
3	56	Warm - Total
3	57	Warm - Joint
3	58	Punitive - Total
3	59	Punitive - Joint
3	60	Confidence - Total
3	61	Confidence - Joint
3	62	Supportive - Total
3	63	Supportive - Joint
3	64	Stimulating - Total
3	65	Stimulating - Joint
3	66	Achievement Oriented - Total
3	67	Achievement Oriented - Joint
3	68	Physical Contact - Total
3	69	Physical Contact - Joint
3	72	Intellectual Effectiveness - Total
3	73	Intellectual Effectiveness - Joint
3	74	Disparaging - Total
3	75	Disparaging - Joint

Card No.	Variable Number	Name of Variable
Identification		
3	76	Friendly vs Hostile - Total (Class)
3	77	Friendly vs Hostile - Joint (Class)
3	78	Academic - Total (Class)
3	79	Academic - Joint (Class)
3	80	Inhibited - Total (Class)
3	81	Inhibited - Joint (Class)
Factor No. 1		
3	82	Total
3	83	Joint
Observation Categories (Stannines)		
3	84	Answers Raised Hand - Total
3	85	Praise and Encouragement - Total
3	86	Non-verbal Affiliation - Total
3	87	Minimal Reinforcement - Total
3	88	Asking Questions - Total
3	89	Problem Structuring - Total
3	90	Academic Control - Total
3	91	Personal Control - Total
3	92	Moralizing - Total
3	93	Hostility and Reprimands - Total
3	94	Ignores Child - Total
Identification		
3	Card No.	

Card No.	Variable Number	Name of Variable
Identification		
4		Grade
4		School
4		Teacher No.
Observation Categories (Percentages)		
4	95	Answers Raised Hand - Total
4	96	Praise and Encouragement - Total
4	97	Non-verbal Affiliation - Total
4	98	Minimal Reinforcement - Total
4	99	Asking Questions - Total
4	100	Problem Structuring - Total
4	101	Academic Control - Total
4	102	Personal Control - Total
4	103	Moralizing - Total
4	104	Hostility and Reprimands - Total
4	105	Ignores Child - Total
Observation Ratings		
4	106	Permissive vs Controlling - Total
4	107	Dull vs Stimulating - Total
4	108	Disparaging vs Less Disparaging - Total
4	109	Supportive vs Less Supportive - Total
4	110	Anxious vs Confident - Total
4	111	Aloof vs Affiliative - Total
4	112	Intent vs Relaxed - Total
4	113	Chaotic vs Orderly - Total (Class)
4	114	Happy vs Unhappy - Total (Class

Card No.	Variable Number	Name of Variable
Identification		
4	115	Dependent vs Independent - Total (Class)
4	116	Smiling vs Sour - Total (Class)
Identification		
4		Card No.
Identification		
5		Grade
5		School
5		Teacher No.
Fall Reading		
5	117	Vocabulary Total
Spring Reading		
5	118	Vocabulary Total
5	119	Comprehension Total
Spring Arithmetic		
5	120	Total Score
Fall Questionnaire		
5	121	N - Achievement
5	122	N - Affiliation
5	123	N - Dependency
5	124	Anxiety
Spring Questionnaire		
5	125	TAS
5	126	P. Affiliation
5	127	P. Control

Card No.	Variable Number	Name of Variable
Identification		
5	128	P. Achievement
5	129	Liking for School
5	130	N - Achievement
5	131	N - Affiliation
5	132	N - Dependency
5	133	Anxiety
5	134	Total Anxiety
Creativity		
5	135	Fall Total
5	136	Spring Total
Creativity Rating		
5	137	Fall
5	138	Spring
Sociometric		
5	139	Received
5	140	Given
Identification		
5		Card No.

THIRD GRADE TEACHER MATRIX

Card No.	Variable Number	Name of Variable
Identification		
1		Grade
1		School
1		Teacher No.
Teaching Preference Schedule		
1	1	N - Achievement
1	2	N - Affiliation
1	3	N - Recognition
1	4	N - Control (Constant = + 40)
1	5	GS1 - Instrumental Rewards
1	6	GS2 - Status Striving
1	7	GS3 - Childs' Affection
1	8	GS4 - Childs' Autonomy
1	9	GS5 - Rebellious Motives
1	10	GS6 - Vicarious Motives
1	11	GS7 - Obsessive
1	12	GS8 - Dependency
1	13	GS9 - Exhibitionism
1	14	GS10 - Dominance
Personal Preferences for Educational Objectives		
1	15	N - Achievement
1	16	N - Affiliation

Card No.	Variable Number	Name of Variable
Identification		
1	17	N - Recognition
1	18	N - Control
Edwards Personal Preference Schedule		
1	19	Achievement
1	20	Order
1	21	Autonomy
1	22	Intraception
1	23	Dominance
1	24	Nurturance
1	25	Endurance
1	26	Aggression
1	27	Deference
1	28	Exhibition
1	29	Affiliation
1	30	Succorance
1	31	Abasement
1	32	Change
1	33	Heterosexuality
1	34	Cons
Identification		
1		Card Number
Identification		
2		Grade
2		School

Card No.	Variable Number	Name of Variable
Identification		
2		Teacher
Situations Test		
2	35	N - Achievement
2	36	N - Affiliation
2	37	N - Recognition
2	38	N - Control
Biographical Data		
2	39	Age
2	40	Years Taught
2	41	College
2	42	Degree
2	43	Hours Over Graduation
2	44	Year Degree Received
2	45	Major
Problem Situation Inventory		
2	46	Control
House - Tree - Person		
2	47	Control
2	48	Affiliation
2	49	Achievement
2	50	Recognition
2	51	Ego
Problem Situation Inventory		
2	52	Item (Affiliation)

Card No.	Variable Number	Name of Variable
Identification		
2	53	Q (Affiliation)
Sociometrics		
2	54	Aggression - Student
2	55	Dependency - Student
2	56	Achievement - Student
2	57	Affiliation - Student
2	58	Anxiety - Student
2	59	Aggression - Teacher
2	60	Dependency - Teacher
2	61	Achievement - Teacher
2	62	Affiliation - Teacher
2	63	Anxiety - Teacher
Identification		
2		Card No.
Identification		
3		Grade
3		School
3		Teacher No.
Q-Sorts		
3	64	Controlling - Total
3	65	Controlling - Joint
3	66	Warm - Total
3	67	Warm - Joint
3	68	Punitive - Total

Card No.	Variable Number	Name of Variable
Identification		
3	69	Punitive - Joint
3	70	Confidence - Total
3	71	Confidence - Joint
3	72	Supportive - Total
3	73	Supportive - Joint
3	74	Stimulating - Total
3	75	Stimulating - Joint
3	76	Achievement Oriented - Total
3	77	Achievement Oriented - Joint
3	78	Physical Contact - Total
3	79	Physical Contact - Joint
3	82	Intellectual Effectiveness - Total
3	83	Intellectual Effectiveness - Joint
3	84	Disparaging - Total
3	85	Disparaging - Joint
3	86	Friendly vs Hostile - Total (Class)
3	87	Friendly vs Hostile - Joint (Class)
3	88	Academic - Total (Class)
3	89	Academic - Joint (Class)
3	90	Inhibited - Total (Class)
3	91	Inhibited - Joint (Class)

Card No.	Variable Number	Name of Variable
Identification		
Factor No. 1		
3	92	Total
3	93	Joint
Observation Categories (Stannines)		
3	94	Answers Raised Hand - Total
3	95	Praise and Encouragement - Total
3	96	Non-verbal Affiliation - Total
3	97	Minimal Reinforcement - Total
3	98	Asking Questions - Total
3	99	Problem Structuring - Total
3	100	Academic Control - Total
3	101	Personal Control - Total
3	102	Moralizing - Total
3	103	Hostility and Reprimands - Total
3	104	Ignores Child - Total
Identification		
3		Card No.
Identification		
4		Grade
4		School
4		Teacher No.
Observation Categories (Percentages)		
4	105	Answers Raised Hand - Total
4	106	Praise and Encouragement - Total

Card No.	Variable Number	Name of Variable
4	107	Non-verbal Affiliation - Total
4	108	Minimal Reinforcement - Total
4	109	Asking Questions - Total
4	110	Problem Structuring - Total
4	111	Academic Control - Total
4	112	Personal Control - Total
4	113	Moralizing - Total
4	114	Hostility and Reprimands - Total
4	115	Ignores Child - Total
Observation Ratings		
4	116	Permissive vs Controlling - Total
4	117	Dull vs Stimulating - Total
4	118	Disparaging vs Less Disparaging - Total
4	119	Supportive vs Less Supportive - Total
4	120	Anxious vs Confident - Total
4	121	Aloof vs Affiliative - Total
4	122	Intent vs Relaxed - Total
4	123	Chaotic vs Orderly - Total (Class)
4	124	Happy vs Unhappy - Total (Class)
4	125	Dependent vs Independent - Total (Class)
4	126	Smiling vs Sour - Total (Class)
Identification		
4		Card No.
Identification		
5		Grade

Card No.	Variable Number	Name of Variable
Identification		
5		School
5		Teacher No.
Fall Reading		
5	127	Vocabulary Total
Spring Reading		
5	128	Vocabulary - Total
5	129	Comprehension - Total
Spring Arithmetic		
5	130	Total Score
Fall Questionnaire		
5	131	N - Achievement
5	132	N - Affiliation
5	133	N - Dependency
5	134	Anxiety
Spring Questionnaire		
5	135	TAS
5	136	P. Affiliation
5	137	P. Control
5	138	P. Achievement
5	139	Liking for School
5	140	N - Achievement
5	141	N - Affiliation
5	142	N - Dependency
5	143	Anxiety

Card No.	Variable Number	Name of Variable
Identification		
5	144	Total Anxiety
Creativity		
5	145	Fall Total
5	146	Spring Total
Creativity Rating		
5	147	Fall
5	148	Spring
Barron Welch		
5	149	Fall
5	150	Spring
Identification		
5		Card No.

APPENDIX C

REGRESSION EQUATIONS USED IN OBTAINING PUPIL 'RECRESSED GAIN' SCORES

Predictor Variable X	Predicted Variable Y	Prediction Equation	r_{xy}	r_{xy}*
	GRADE I			
Fall Read. Vocab.	Spring Read. Vocab.	$\widetilde{Y} = .55X + 30.2$	.48	.57
Fall Circles Orig.	Spring Circles Orig.	$\widetilde{Y} = .42X + 13.2$	.32	
	GRADE III			
Fall Read. Vocab.	Spring Read. Vocab.	$\widetilde{Y} = .67X + 12.3$	.67	.72
Fall Read. Vocab.	Spring Read. Comp.	$\widetilde{Y} = .92X + 11.1$	.69	.68
Fall Read. Vocab.	Spring Arithmetic	$\widetilde{Y} = .30X + 14.8$	.40	.37
Fall Anxiety	Spring Anxiety	$\widetilde{Y} = .30X + 2.0$	.38	
Fall Barron Welch	Spring Barron Welch	$\widetilde{Y} = .78X + 4.2$	.68	
Fall Circles--Rating	Spring Circles--Rat.	$\widetilde{Y} = .45X + 1.5$	.41	

*Correlations obtained for the same variable in our prior study--two years before. In the earlier study Fall Reading Vocabulary correlated .53 and .52 with Spring Reading Comprehension and Arithmetic in Grade I. In the present study, thes values dropped to .25 and .29.

SELECTED REFERENCES

Barr, A. S. The Measurement and Prediction of Teacher Efficiency: A Summary of Investigations, J. exp. Educ., 1948, 16, 203-83.

Barr, A. S., Bechdolt, B. V., Coxe, W. W., Gage, N. L., Orleans, J. S., Remmers, H. H., and Ryans, D. G. Report of the Committee on Criteria of Teacher Effectiveness, Rev. educ. Res., 1952, 22, 238-263.

Cogan, M. L. The Behavior of Teachers and the Productive Behavior of Their Pupils, J. exp. Educ., 1958, 27, 89-124.

Damrin, D. The Russell Sage Social Relations Test. A Measure of Group Problem Solving Skills in Elementary School Children. Proceedings of the Invitational Conference on Testing Problems, 1954, 75-84.

Flanders, N. A. Teacher Influence, Pupil Attitudes and Achievement. Pre-publication manuscript of proposed research monograph for U. S. Office of Education, Cooperative Research Branch, Research Project No. 397, 1962.

Heil, L. M., Powell, M., and Feifer, I. Characteristics of Teacher Behavior Related To the Achievement of Children in Several Elementary Grades. Final Report, U. S. Office of Education Contract No. SAE 7285, 1960.

Jayne, C. D. A Study of the Relationship Between Teaching Procedures and Educational Outcomes, J. exp. Educ., 1945, 14, 101-134.

Medley, D. M. and Klein, A. A. Measuring Classroom Behavior with a Pupil Reaction Inventory, Elem, sch. J., 1957, 57, 315-319.

Mitzel, H. E. and Medley, D. M. Pupil Growth in Reading: An Index of Effective Teaching, J. educ. Psychol., 1957, 48, 227-239.

Marsh, J. E., Burgess, F. F. and Sinilk, P. W. Student Achievement as a Measure of Instructor Effectiveness, J. educ. Psychol., 1956, 47, 79-88.

Ryans, D. G. Characteristics of Teachers. American Council on Education, Washington, 1960.

Ryans, D. G. Prediction of Teaching Effectiveness. In Harris, C. W. (Ed.) Encyclopedia of Educational Research, 1960.

Stern, G. G. and Masling, J. M. Unconscious Factor in Career Motivation for Teaching. Final Report, U. S. Office of Education Project, No. SAE 6459, 1958.

Sarason, S. B., et al. Anxiety in Elementary School Children. New York: Wiley and Sons, 1960.

Travers, R. M. W., Wallen, N. E., Reid, I. E., and Wodtke, K. H. Measured Needs of Teachers and Their Behavior in the Classroom. Final Report, U. S. Office of Education, Contract No. 444(8029), 1961.

Wallen, N. E. and Travers, R. M. W. Analysis and Investigation of Teaching Methodologies. In Gage, N. L. (Ed.) Handbook of Research on Teaching. American Educational Research Association, Rand-McNally, 1963.

Wallen, N. E., Travers, R. M. W., Reid, I. E. and Wodtke, K. H. Relationships Between Teacher Needs and Teacher Behavior in the Classroom, J. educ. Psychol., 1963, 54, 23-32.

Wallen, N. E. and Wodtke, K. H. Relationships Between Teacher Characteristics and Student Behavior: Part I. Final Report, U. S Office of Education, Contract No. 2-10-013, 1963.

Wallen, N. E. Relationships Between Teacher Characteristics and Student Behavior: Part II. Final Report, U. S. Office of Education, Contract No. 4-10-034, 1964.

Yamamoto, K. Revised Scoring Manual for Tests of Creative Thinking. Bureau of Educational Research, University of Minnesota, Minneapolis, 1962.